金融情報学研究会(第24回)

日時　2020 年 3 月 14 日(土)・15 日(日)

会場　成蹊大学 ６号館４階４０１教室

SIG-FIN
JSAI Special Interest Group on Financial Informatics

人工知能学会
金融情報学研究会

第24回 人工知能学会 金融情報学研究会(SIG-FIN)

2020 年 3 月 14 日(土)　成蹊大学 6号館

01. 時系列モデルを用いたマルチアセット市場における統計的裁定戦略 …… 1
今井崇公，中川慧(野村アセットマネジメント)

02. 遺伝的プログラミングを用いたテクニカル指標による金融取引の戦略木構築 …… 7
加藤旺樹，穴田一(東京都市大学)

03. 人工市場を用いた負の売買手数料(リベート)が株式市場に与える影響の調査 …… 12
星野真広(神奈川工科大学)，水田孝信(スパークスアセットマネジメント)，八木勲(神奈川工科大学)

04. 裁定取引へのリベートは流動性を向上させるか？-人工市場によるシミュレーション分析- …… 20
水田孝信(スパークスアセットマネジメント)

05. 人工市場を用いた株式貸借市場の流動性の変化が売買市場に与える影響の分析 …… 28
宮崎文吾，中川慧(野村アセットマネジメント)

06. 人工市場を用いたレバレッジド ETF がザラ場市場に与える影響分析 …… 35
丸山隼矢(神奈川工科大学)，水田孝信(スパークスアセットマネジメント)，八木勲(神奈川工科大学)

07. DSGE によるニューラルネットの正則化と経済予想 …… 42
塩野剛志(クレディ・スイス証券)

08. 銘柄属性と市場情報を用いた深層学習株式リターン予測モデルの予測根拠の分析 …… 50
小寺俊哉，佐藤史仁(日興リサーチセンター)，坂地泰紀，和泉潔(東京大学)

09. 変分自己符号化器を用いたネットワーク埋め込みと金融ネットワークへの応用 …… 58
川上雄大(神戸大学)，江口浩二(広島大学)

10. 自己符号化器を用いた国債イールドカーブのファクターモデルの構築 …… 66
水門善之(東京大学，野村證券)，坂地泰紀，和泉潔，島田尚，松島裕康(東京大学)

11. 日銀「主な意見」の発言者分類モデルの作成と政策変更予測への応用 …… 70
末廣徹，木村柚里，稲垣真太郎(みずほ証券)

12. ニュース記事に基づく景気指標 S-APIR の開発 …… 76
関和広(甲南大学)，生田祐介(大阪産業大学)，松林洋一(神戸大学)

13. 景気の先行き判断の「先行き」の解釈における差異の識別: …… 84
PU 学習とテキストマイニングの応用
加藤真大(東京大学)

14. アナリストレポートにおけるキーワード関連文の抽出と景況感推移観測への応用 …… 92
高山将丈，小澤誠一(神戸大学)，廣瀬勇秀，飯塚正昭，渡辺一男，
逸見龍太(三井住友 DS アセットマネジメント)

15. 接触履歴を用いた地方景況感と業種間構造の分析 …… 98
坂地泰紀，蔵本涼太，和泉潔，松島裕康，島田尚(東京大学)，砂川恵太(沖縄銀行)

16. 外国為替市場におけるポジション管理戦略分類に基づく将来価格形成メカニズム解明 …… 103
末重拓己(東京工業大学), Didier Sornette(スイス連邦工科大学チューリッヒ校),
高安秀樹(Sony CSL)，高安美佐子(東京工業大学)

17. リターンに基づくクラスタリングを用いた投資信託スタイル分析 …… 108
副島聡一郎(名古屋工業大学)，松井藤五郎(中部大学)，犬塚信博，武藤敦子，
森山甲一(名古屋工業大学)

18. t 過程潜在変数モデルによるポートフォリオ生成 …… 112
内山祐介(MAZIN)，中川慧(野村アセットマネジメント)

2020 年 3 月 15 日(日)　成蹊大学 6号館

19. 人工市場を用いた高頻度取引が金融市場流動性に与える影響の調査 …… 117
益田裕司(神奈川工科大学), 水田孝信(スパークスアセットマネジメント), 八木勲(神奈川工科大学)

20. 強化学習を用いた期待効用ベースヘッジ手法 …… 125
上田翼(三井住友 DS アセットマネジメント)

21. 強化学習による高頻度取引戦略の構築 …… 129
小林弘幸, 和泉潔, 松島裕康, 坂地泰紀(東京大学)

22. 機械学習による金融時系列モデルのパラメータ推定 …… 134
高石哲弥(広島経済大学)

23. 会社四季報を用いた理論株価の推定 : Kaggle を参考に …… 138
吉川満(大和大学)

24. 機械学習によるテクニカル分析の影響の調査 …… 144
片寄諒亮, 吉岡真治(北海道大学)

25. 多相アテンション LSTM を用いた財務時系列データの予測 …… 149
吉見脩平(神戸大学), 江口浩二(広島大学), 金京拓司, 羽森茂之(神戸大学)

26. 経済金融時系列における歪みの検出 …… 157
土屋太助

27. ニュース重要単語の機械学習によるアクティブ運用 …… 160
三好勝博, 細木唯以(茨城大学), 江口潤一(大和投資信託), 鈴木智也(茨城大学, 大和投資信託)

28. 経済因果チェーン検索システムの構築と応用 …… 167
和泉潔, 坂地泰紀(東京大学)

29. 経済因果チェーンを用いたリードラグ効果の実証分析 …… 171
中川慧, 指田晋吾(野村アセットマネジメント), 坂地泰紀, 和泉潔(東京大学)

30. トピック埋め込み回帰モデルを用いた株価予測 …… 177
許蔚然(神戸大学), 江口浩二(広島大学)

31. ニューステキストを用いた ESG ファクター運用 …… 183
秋山祥伍(茨城大学), 江口潤一(大和投資信託), 鈴木智也(茨城大学, 大和投資信託)

32. グラフエンベディングを活用した潜在取引関係予測 …… 187
藤塚理史(日本総合研究所), 工藤剛(三井住友銀行)

33. 日本株式市場における daytime と overnight return の間の VC 相関解析 …… 193
落合友四郎(大妻女子大学), Jose Nacher(東邦大学)

34. B2B 市場における企業ブランドと ROE の関連性 …… 198
真鍋友則(Sansan), 中川慧(野村アセットマネジメント)

35. GCN による取引関係グラフからの企業の特徴量抽出 …… 206
森正和, 與五澤守(日本総合研究所), 工藤剛(三井住友銀行)

36. 株式・債券・為替ごとの金融極性辞書の自動構築 …… 211
伊藤友助, 酒井浩之(成蹊大学), 北島良三(東京工芸大学), 末廣徹, 稲垣真太郎,
木村柚里(みずほ証券)

37. 新興市場を対象とした市況情報の抽出 …… 219
神田裕輝, 高野海斗, 酒井浩之(成蹊大学), 北島良三(東京工芸大学),
中川慧(野村アセットマネジメント)

38. テキストマイニングを利用したテーマに関連する上場企業検索ツールの開発 …… 226
平野正徳, 坂地泰紀(東京大学), 木村笙子(大和投資信託), 和泉潔, 松島裕康(東京大学),
長尾慎太郎(大和投資信託), 加藤惇雄(大和総研)

39. 文脈を考慮した決算短信からの業績要因抽出 …… 234
加藤悠太, 酒井浩之(成蹊大学)

時系列モデルを用いた マルチアセット市場における統計的裁定戦略

Statistical Arbitrage Strategy in Multi-Asset Market using Time Series Analysis

今井　崇公[1*]　中川　慧[1]
Takahiro Imai[1]　Kei Nakagawa[1]

[1] 野村アセットマネジメント株式会社
[1] Nomura Asset Management Co.,Ltd.

Abstract: The statistical arbitrage strategy is one of the most traditional investment strategies. Many theoretical and empirical studies have been conducted for a long time. Almost all of the statistical arbitrage strategies focus on the price difference (spread) between two similar assets in same asset class and exploit the mean reversion of spreads, i.e., pairs trading. In this study, we extend the strategy to multiple assets in the multi-asset market. Concretely, we derive a mean-reverting portfolio with time series model. Finally, we perform an empirical analysis in multi-asset market and show the effectiveness of our strategy.

1　はじめに

資産間の価格差（スプレッド）の平均回帰性に着目した研究は古くから数多く行われており、その性質を利用したペア・トレード戦略は現在でも個人投資家から機関投資家まで幅広い層に利用されている [1]。[2] によると、ペア・トレード戦略を行うにあたってのアプローチは多岐にわたり、銘柄間距離、共和分を代表とする時系列モデル、確率制御を利用するものなどが存在する。一方で、ペア・トレード戦略の多くが類似した銘柄間のスプレッドに着目した戦略であることから、ほとんどの先行研究は単一資産内の銘柄をユニバースとしている。

そこで本研究では、単一資産ではなくマルチアセット市場を対象に時系列モデルを用いて平均回帰性を有するポートフォリオを構築することで、異なる資産クラスの銘柄をロング、ショートするペア・トレード戦略を提案する。マルチアセット市場を対象とすることで、単一資産に留まらないグローバルなアセットクラス間の連動性のずれを裁定機会として得ることが狙いである。具体的には、異なる観点から定義された平均回帰性を定量化する基準を複数用意し、それら複数の基準で見て「良い」ポートフォリオを構築する。その際に、Polynomial Goal Programming(PGP) と呼ばれる多目的最適化問題の手法を用いることで、平均回帰性の定量基準を恣意性なく組み合わせる手法を提案する。そして当該ポートフォリオが平均水準から一定程度乖離したタイミングでポジションをとって、平均回帰する方向にベットする。

本稿の構成は次の通りである。2 章では、関連する時系列モデルを用いた平均回帰ポートフォリオの先行研究について簡単に述べる。3 章では、平均回帰性の「良さ」を示す基準を複数導入すると共に、それらを統合し、最適な保有比率を求める手法である PGP について説明する。そして 4 章において、本研究で対象とするマルチアセット市場におけるペア・トレード戦略を記述し、5 章で実際の金融市場のデータを用いた実証分析によってその有効性を評価する。最後に、結論を述べる。

2　先行研究

平均回帰性の定量基準は様々な形で提案されている。有名な定量基準としては、時系列が分散で測ってどれだけホワイトノイズに近いかを示す Predictability[3, 4]、時系列が相関で測ってどれだけホワイトノイズに近いかを示す Portmanteau Statistics[5]、時系列がある単位時間区間において平均水準に何回回帰したかを示す Crossing Statistics[6] などが存在する。

*連絡先：野村アセットマネジメント株式会社
〒 103-0027　東京都中央区日本橋 1-11-1
E-mail: t-imai@nomura-am.co.jp
本稿の内容は筆者らが所属する組織を代表するものではなく、本稿の全ての誤りは、筆者らの責に属するものである。

これらの指標を用いたペア・トレード戦略として、[7] では米国株式のインプライドボラティリティを、[8] では米国株式を投資対象に Predictability、Portmanteau Statistics、Crossing Statistics をそれぞれ単独に用いて平均回帰ポートフォリオを構築している。 これらの研究は単体の平均回帰性の定量基準の有効性を評価できているという点で有用ではあるが、多角的な視点で平均回帰ポートフォリオを構築できてはいない。ここで複数の平均回帰性の基準を組み合わせる方法として、サブ目的関数を結合した多目的最適化問題を解く方法が考えられるが、最適化の結果に大きく関わる結合係数の決定方法について恣意性が残るという課題がある。このような多目的最適化問題のハイパーパラメータ設定を容易にする方法として、PGP と呼ばれる手法が高次モーメントを含めたポートフォリオ最適化問題においてしばしば用いられている [9]。

3 平均回帰ポートフォリオの構築手法

N 個の資産が存在する時、その時点 t における対数価格を $\boldsymbol{y_t} = \{y_{1,t}, ..., y_{N,t}\}$ とする。各資産のウェイトベクトルを $\boldsymbol{w} = \{w_1, ..., w_N\}$ とすると、ポートフォリオは

$$z_t = \boldsymbol{w}^\top \boldsymbol{y_t} \tag{1}$$

と書け、ポートフォリオの対数収益率 r_t は、

$$r_t = z_t - z_{t-1} = \boldsymbol{w}^\top (\boldsymbol{y_t} - \boldsymbol{y_{t-1}}) \tag{2}$$

と表せる。本研究における問題設定は、式 (1) のポートフォリオ z_t が平均回帰するようなウェイト $\boldsymbol{w}$ を求めることである。すなわち、複数資産間においてできるだけ平均回帰するようなウェイトを算出することが目的である。

3.1 平均回帰の基準

ここでは、ポートフォリオ z_t の平均回帰の良さを定量的に示す基準を複数導入する。本研究では、(1)Predictability、(2)Portmanteau Statistics、(3)Crossing Statistics の 3 つの観点で平均回帰性を定量化する。はじめに、系列 $\boldsymbol{y_t}$ に対する次数 (ラグ)$i \in \mathbb{N}$ の自己共分散行列を次の通り定義する。

$$\begin{aligned} \boldsymbol{M_i} &:= \mathrm{Cov}\left(\boldsymbol{y_t}, \boldsymbol{y_{t+i}}\right) \\ &= \mathbb{E}\left[\left(\boldsymbol{y_t} - \mathbb{E}[\boldsymbol{y_t}]\right)\left(\boldsymbol{y_{t+i}} - \mathbb{E}[\boldsymbol{y_{t+i}}]\right)^\top\right] \end{aligned} \tag{3}$$

式 (3) から明らかであるが、$i = 0$ の $\boldsymbol{M_0}$ は通常の共分散行列を表す。

3.1.1 Predictability

この基準は時系列が分散で測ってどれだけホワイトノイズに近いかという観点から平均回帰性を定量的に表す。定常な以下の時系列を考える。

$$y_t = \hat{y}_{t-1} + \varepsilon_t \tag{4}$$

ここで、$\hat{y}_{t-1}$ は時刻 $t-1$ までの情報に基づく予測値を表し、最も単純な例として $\hat{y}_{t-1}$ を αy_{t-1} とする AR(1) モデルがある。ε_t は $\hat{y}_{t-1}$ とは独立の分散 σ_ε^2 のホワイトノイズを表す。式 (4) の分散をとると、$\sigma_y^2 = \sigma_{\hat{y}}^2 + \sigma_\varepsilon^2$ となる。ここで、σ_y^2 は y_t の分散、$\sigma_{\hat{y}}^2$ は $\hat{y}_{t-1}$ の分散をそれぞれ表す。

Predictability は、

$$predictability = \frac{\sigma_{\hat{y}}^2}{\sigma_y^2} \tag{5}$$

と定義される。

Predictability の定義から、この値が小さいと、ホワイトノイズに近いことを意味し、平均回帰しやすい。一方で値が大きいと、$t-1$ までの情報で予測しやすいことを意味する。

$\hat{y}_{t-1}$ を次の VAR(1) モデルでモデル化できるとする。なお、[10] から一般に VARMA(p, q) モデルは VAR(1) モデルで表現可能であるため、以下の議論は VARMA(p, q) モデルに対しても同様に可能である。

$$\boldsymbol{y_t} = \boldsymbol{A y_{t-1}} + \boldsymbol{e_t} \tag{6}$$

ここで $\boldsymbol{e_t}$ はホワイトノイズである。

式 (6) の VAR(1) モデルに $\boldsymbol{w}^\top$ をかけると、$\boldsymbol{w}^\top \boldsymbol{y_t} = \boldsymbol{w}^\top \boldsymbol{A y_{t-1}} + \boldsymbol{w}^\top \boldsymbol{e_t}$ となる。この分散を取ると、左辺は $\boldsymbol{w}^\top \boldsymbol{M_0 w}$ となり、右辺第一項 $\boldsymbol{w}^\top \boldsymbol{A y_{t-1}}$ は $\boldsymbol{w}^\top \boldsymbol{A M_0 A}^\top \boldsymbol{w}$ となる。さらに VAR(1) モデルの性質から $\boldsymbol{A} = \boldsymbol{M_1}^\top \boldsymbol{M_0}^{-1}$ であるので、これを代入して結局、右辺第一項は $\boldsymbol{w}^\top \boldsymbol{A M_0 A}^\top \boldsymbol{w} = \boldsymbol{w}^\top \boldsymbol{M_1}^\top \boldsymbol{M_0}^{-1} \boldsymbol{M_0} \left(\boldsymbol{M_1}^\top \boldsymbol{M_0}^{-1}\right)^\top \boldsymbol{w} = \boldsymbol{w}^\top \boldsymbol{M_1}^\top \boldsymbol{M_0}^{-1} \boldsymbol{M_1 w}$ となる。したがって、VAR(1) モデルの Predictability は、

$$predictability(\boldsymbol{w}) = \frac{\boldsymbol{w}^\top \boldsymbol{M_1}^\top \boldsymbol{M_0}^{-1} \boldsymbol{M_1 w}}{\boldsymbol{w}^\top \boldsymbol{M_0 w}} \tag{7}$$

となる。

3.1.2 Portmanteau Statistics

この基準は時系列が自己相関で測ってどれだけホワイトノイズに近いかという観点から平均回帰性を定量的に表す。

定常なラグ p を持つ y_t に対して、

$$y_t = \hat{y}_{t-1} + ...\hat{y}_{t-p} + \varepsilon_t \tag{8}$$

Portmanteau Statistics は

$$portmanteau = \sum_{i=1}^{p} \rho_i^2 \tag{9}$$

とかける。ここで ρ_i は、ラグ i の自己相関係数であり、$\mathbb{E}[z_t z_{t+i}]/\mathbb{E}[z_t^2]$ と定義される。式 (10) の定義から $portmanteau \geq 0$ であることがわかり、またホワイトノイズはどのラグ p ととっても自己相関を持たないので、$portmanteau$ が 0 に近ければ、自己相関で測ってホワイトノイズに近いと言える。したがってこれを最小化するように $\boldsymbol{w}$ をとると平均回帰しやすい。

$z_t = \boldsymbol{w}^\top \boldsymbol{y_t}$ の時、自己相関係数は $\rho_i = \mathbb{E}\left[z_t z_{t+i}\right]/\mathbb{E}\left[z_t^2\right]$ より、平均回帰ポートフォリオの portmanteau は次のように書ける。

$$portmanteau(\boldsymbol{w}) = \sum_{i=1}^{p} \left(\frac{\boldsymbol{w}^\top \boldsymbol{M_i} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{M_0} \boldsymbol{w}} \right)^2 \tag{10}$$

3.1.3 Crossing Statistics

この基準は時系列がある時間区間 T において平均水準に何回回帰したかをカウントし、その数で平均回帰性を定量的に表す。

Crossing statistics は、定常時系列 y_t に対して、以下の通り定義される。

$$crossing = \frac{1}{T-1} \sum_{t=2}^{T} \mathbf{1}_E\left(y_t\right) \tag{11}$$

ここで、$\mathbf{1}_E\left(y_t\right)$ は、事象 $E = \{y_t y_{t-1} \leq 0\}$ を満たす時に 1、満たさない場合には 0 を返す指示関数である。

定常な正規ノイズを持つ時系列に対しては、

$$crossing = \frac{1}{\pi} \arccos\left(\rho_1\right) \tag{12}$$

で与えられることが知られている。この式から、定常な正規ノイズを持つ時系列では、平均水準に多く回帰させるためには、ρ_1 を小さくすることが必要である。

以上から $z_t = \boldsymbol{w}^\top \boldsymbol{y_t}$ の時、平均回帰ポートフォリオの Crossing statistics は 1 次の自己相関の定義から次のように書ける。

$$crossing(\boldsymbol{w}) = \frac{\boldsymbol{w}^\top \boldsymbol{M_1} \boldsymbol{w}}{\boldsymbol{w}^\top \boldsymbol{M_0} \boldsymbol{w}} \tag{13}$$

3.2 Polynomial Goal Programming による定式化

ここでは、前節で導入した 3 つの平均回帰性の定量基準を PGP により統合し、最適な保有比率を決定する。この手法は、高次モーメントを含めたポートフォリオの多目的最適化問題において導入された手法で、複数のサブ目的関数を規格化する点と、それらのサブ目的関数に投資家の選好を反映できる点が利点である。

PGP では、まずサブ目的関数である式 (7)、式 (10)、式 (13) の最小化問題を独立に解き、各々の最小値 $pred^*$、$port^*$、$cross^*$ を求める。次に、それらを用いて

$$\begin{aligned} \arg\min_{\boldsymbol{w}} \quad & \left|\frac{d_1}{pred^*}\right|^{\lambda_1} + \left|\frac{d_2}{port^*}\right|^{\lambda_2} + \left|\frac{d_3}{cross^*}\right|^{\lambda_3} \\ \text{where} \quad & d_1 = predictability(\boldsymbol{w}) - pred^* \\ & d_2 = portmanteau(\boldsymbol{w}) - port^* \\ & d_3 = crossing(\boldsymbol{w}) - cross^* \end{aligned} \tag{14}$$

を解くことで、最適な $\boldsymbol{w}$ を求める。ここで λ_1、λ_2、λ_3 はそれぞれ Predictability、Portmanteau Statistics、Crossing Statistics の選好度を表す変数である。

3 つの平均回帰性に関する定量基準と PGP による統合について、概念図を図 1 に示す。

4 提案手法

本研究では、前章で導入した定量基準を基にポートフォリオを構築し、移動平均からの乖離が平均回帰する性質を利用したペア・トレード戦略を提案する。

Step 1：

マルチアセットを対象に、あらかじめ設定した最大次数以下で VAR(p) モデルの AIC が最小となる次数 p を選択する。この次数を用いて、各平均回帰性に関する定量基準の値を算出する。

Step 2：

(1)Predictability、(2)Portmanteau Statistics、(3)Crossing Statistics、(4)PGP を用いて (1)〜(3) を統合した多目的関数をそれぞれ最小化することで、平均回帰性を有するポートフォリオのウェイトを算出する。

Step 3：

Step 2 で算出したポートフォリオの過去リターンの移動平均からのスプレッドを算出する。スプレッドが平均 ±1 標準偏差以上乖離した時をシグナルとしてポジションを取得する。またロスカットとして、ポジション取得後に平均 ±2 標準偏差以上乖離した場合は、ポジションを解消する。

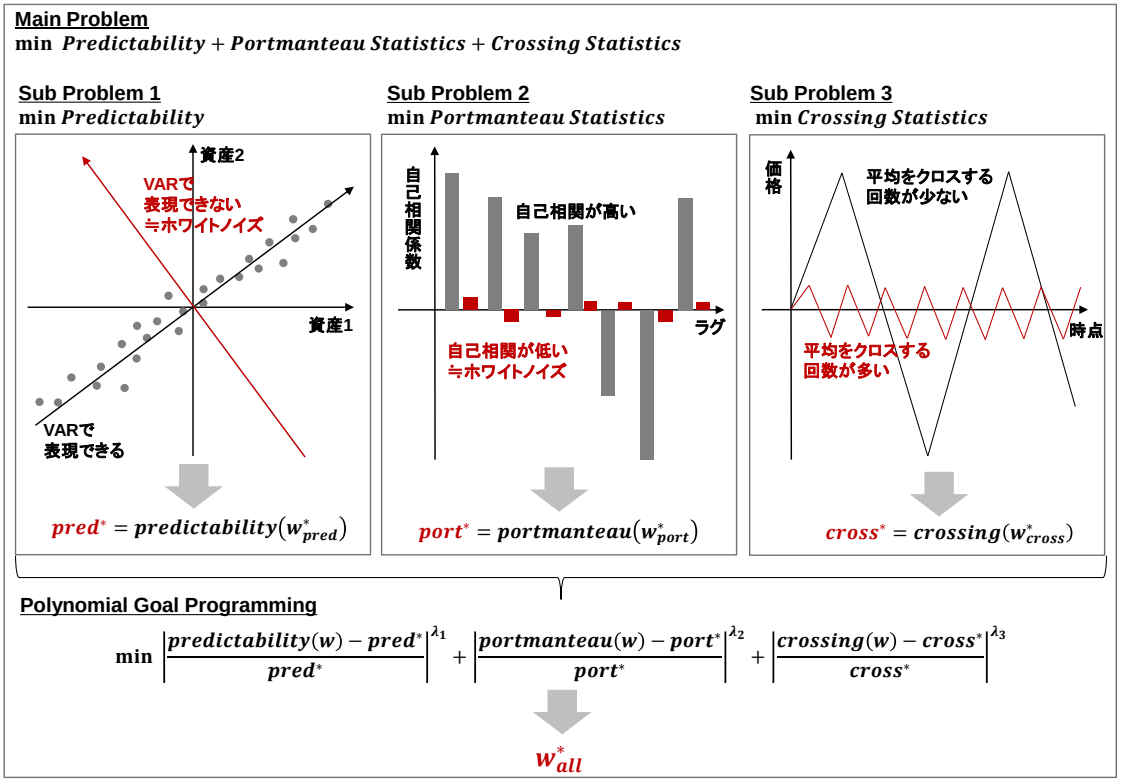


図 1: 3 つの平均回帰性に関する定量基準と PGP による統合

5 実証分析

5.1 データセット

ここではグローバルの先物を対象に分析を行った。投資対象のユニバースを表 1 にまとめた。データは Bloomberg から取得した。検証期間は、2007 年 3 月 23 日から 2019 年 8 月 30 日までとし、すべて金曜引けの週次リターンを用いた。

5.2 分析方法

四半期毎に、最大次数を 5、データ期間を過去 52 週として VAR(p) モデルの次数を決定し、最適なポートフォリオのウェイトを算出した。ここでポートフォリオのレバレッジは、過去 52 週のリターンデータを用いて推定した共分散行列を基に、推定リスクが年率 5%となるように決定した。なお Predictability のみ、PGP で統合することを踏まえ VAR(1) モデルを用いた。仮にポジションを取得していたとしても四半期毎に解消することとした。また、ロスカットを行った場合、当該四半期はポジションを取得しないこととした。PGP における Predictability、Portmanteau Statistics、Crossing Statistics の選好度はすべて 1 とした。ポートフォリオの過去リターンの移動平均は過去 13 週、スプレッドの平均および標準偏差は過去 52 週の値を用いて算出した。以後、ポジション取得から解消までの期間を 1 戦略と呼称する。

5.3 分析結果

各平均回帰性の定量基準を用いて、ポートフォリオを構築した際のパフォーマンス統計量のサマリーを表 2 に、累積リターンの比較を図 2 に示す。

表 2 の各パフォーマンス統計量の定義について簡単に説明する。投資期間のリターン、リスク、リターン/リスクは、それぞれポジションを取得している時のみのポートフォリオのリターン系列を用いて算出しており、リターンおよびリターン/リスクは値が大きいほど、リスクは値が小さいほどパフォーマンスが良いことを示す。戦略の勝率は、ポートフォリオのリターンが正であった戦略数を総戦略数で除した値とし、値が大きいほどパフォーマンスが良いことを示す。戦略の最大ドローダウンは、全戦略のうち最も低いリターンの値を表し、絶対値が小さいほどパフォーマンスが良いことを示す。全期間に占める投資期間の割合は、全期間のうちポジションを取得している期間の割合を表し、1 戦略あたりの平均投資期間は、ポジションを取得してから解消するまでにかかった期間の平均を表す。

表 1: 投資対象資産

	銘柄							
株式先物（計 16）	S&P 500（SP）	NASDAQ（NQ）	カナダ（PT）	英（CF）	仏（CF）	独（GX）	欧州（VG）	スペイン（IB）
	オランダ（EQ）	ノルウェー（OI）	スイス（SM）	日経（NK）	TOPIX（TP）	香港（HI）	豪州（XP）	シンガポール（QZ）
債券先物（計 13）	米国 10 年（TY）	米国 2 年（TU）	米国 5 年（FV）	米国 20 年（US）	豪州 3 年（YM）	豪州 10 年（XM）	カナダ 10 年（CN）	独 2 年（DU）
	独 5 年（OE）	独 10 年（RX）	独 30 年（UB）	英 10 年（G）	日本 10 年（JB）			

表 2: パフォーマンス統計量

	PGP	Predictability	Portmanteau	Crossing
投資期間のリターン（%、年率換算）	**11.9**	9.9	3.1	2.8
投資期間のリスク（%、年率換算）	6.9	8.9	**4.5**	5.6
投資期間のリターン/リスク	**1.73**	1.11	0.70	0.50
戦略の勝率（%）	**59.8**	57.3	50.6	49.4
戦略の最大ドローダウン（%）	-2.5	-4.4	**-1.7**	-2.7
全期間に占める投資期間の割合 （%）	28.0	24.5	30.0	23.8
1 戦略あたりの平均投資期間（週）	1.9	1.8	2.4	2.0

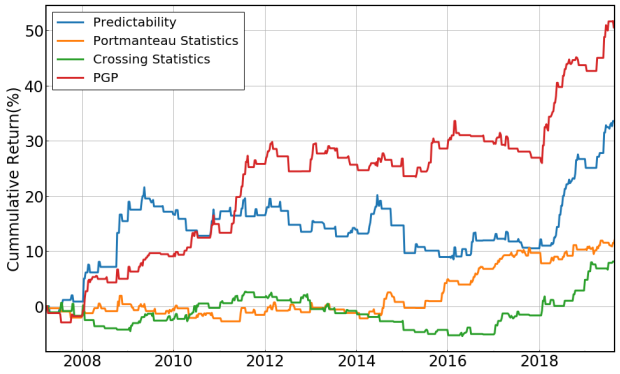


図 2: 累積リターンの比較

表 2 の投資期間のリターン/リスクおよび戦略の勝率では PGP が 3 つの単体手法を優越する結果を示しており、複数の平均回帰性の基準を組み合わせることの有効性を示唆している。図 2 によれば、各単体手法が有効な期間は概ね異なるが、PGP を用いることでこれらをうまく統合できていると言える。

6 まとめ

本研究では、マルチアセット市場を対象に時系列モデルを用いて平均回帰性を有するポートフォリオを構築することで、異なる資産クラスの銘柄をロング、ショートするペア・トレード戦略を提案した。具体的な平均回帰性の定量基準としては、分散の観点で測った Predictability、相関の観点で測った Portmanteau Statistics、平均水準への回帰回数で測った Crossing Statistics、およびこれらを PGP と呼ばれる多目的最適化問題の手法により統合した基準を用いた。実証分析では、PGP を用いた手法において最良のリターン/リスクおよび勝率の結果が得られ、マルチアセット市場のペア・トレード戦略において複数の平均回帰性の定量基準を組み合わせる戦略の有効性が示唆された。

参考文献

[1] Ganapathy Vidyamurthy. *Pairs Trading: quantitative methods and analysis*, volume 217. John Wiley & Sons, 2004.

[2] Christopher Krauss. Statistical arbitrage pairs trading strategies: Review and outlook. *Journal of Economic Surveys*, 31(2):513–545, 2017.

[3] George EP Box and George C Tiao. A canonical analysis of multiple time series. *Biometrika*, 64(2):355–365, 1977.

[4] Ronald Bewley, David Orden, Minxian Yang, and Lance A Fisher. Comparison of box—tiao and johansen canonical estimators of cointegrating vectors in vec (1) models. *Journal of Econometrics*, 64(1-2):3–27, 1994.

[5] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[6] N Donald Ylvisaker. The expected number of zeros of a stationary gaussian process. *The Annals of Mathematical Statistics*, 36(3):1043–1046, 1965.

[7] Marco Cuturi and Alexandre d ' Aspremont. Mean reversion with a variance threshold. In *International Conference on Machine Learning*, pages 271–279, 2013.

[8] Ziping Zhao and Daniel P Palomar. Mean-reverting portfolio with budget constraint. *IEEE Transactions on Signal Processing*, 66(9):2342–2357, 2018.

[9] Mehmet Aksarayl ı and Osman Pala. A polynomial goal programming model for portfolio optimization based on entropy and higher moments. *Expert Systems with Applications*, 94:185–192, 2018.

[10] Helmut Lütkepohl. *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

遺伝的プログラミングを用いた テクニカル指標による金融取引の戦略木構築

Construction of a Financial Transaction Strategy Tree using Genetic Programming by Technical Index

加藤　旺樹[1]　穴田　一[1]

Ohki Kato[1]　Hajime Anada[1]

[1]東京都市大学大学院 総合理工学研究科

[1]Graduate School of Integrative Science and Engineering, Tokyo City University

Abstract: Recently, many researchers have studied stock trading using technical analysis. However, it is necessary to have deep knowledge to use such technical analysis and it is difficult to achieve profitability using this technique. Therefore, using Genetic Programming, we construct an evolutionary model that considers the technical index for devising a profitable trading strategy. Finally, we confirmed the effectiveness of our model using historical data of the stock market.

1. はじめに

近年，テクニカル分析を用いた株式売買に関する研究が精力的に行われている［1］［2］［3］［4］［5］［6］．テクニカル分析を用いた投資では，相場のトレンドや転換点を判断するテクニカル指標を用いることで，過去の値動きのパターンから将来の値動きを予測し，売買を行う．しかし，そのためには専門的な知識を必要とする上，利益を上げにくいという問題がある．

松村らは，テクニカル指標を用いた投資戦略を木構造で戦略木として表現した．そして，それぞれの個体が戦略木を持ち，その戦略木に従い株式売買を行うモデルを構築した．戦略木を構成するノードを個体の遺伝子として遺伝的操作を用いることで，より利益が高くなるように個体を進化させ，各個体が持つ戦略木の特徴を分析した．ところが，その戦略木は上昇トレンドのみでしか稼ぐことができないものとなっていた．

そこで本研究では，松村らの研究で用いられた戦略木を用いて，さまざまなトレンドで利益を生み出す，テクニカル指標を用いた投資戦略の構築を目的とする．

2. 提案手法

それぞれの個体が戦略木を持ち，その戦略木にしたがって取引を行う．その取引結果から個体を評価した値である適応度を求め，戦略木を構成するノードを個体の遺伝子として遺伝的操作を用いることで，より適応度が高くなるように個体を進化させていく．本研究では，既存手法を改良するために，使用するノードの種類の修正，適応度の定義，突然変異の方法，次世代に残す個体の選択方法の変更と進化の過程で戦略木中の使用されないノードの削除を行った．

2.1. テクニカル指標

テクニカル指標は株式の売買タイミングを判断するために使われる指標で，トレンド系，オシレータ系，出来高系の 3 つがある．トレンド系は株価の推移からトレンドを判断する指標，オシレータ系は株価の推移からトレンドの転換点を判断する指標，出来高系は売買が成立した株数の推移からトレンドの転換点を判断する指標である．

2.2. 戦略木

戦略木は，表 1 に示す複数の非終端ノードと終端ノードからツリー状に構成される．図 1 に戦略木の例を示す．

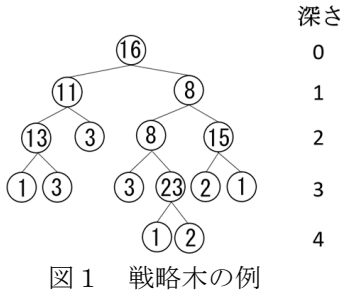


図 1　戦略木の例

それぞれの数字がテクニカル指標や売買行動を表し，各ノードから 1 番上のノードまでのノード間を結ぶ枝の数を各ノードの深さとする．

各個体は戦略木を持ち，戦略木に従って当日の行動を決定する．一番上のノードからテクニカル指標に基づく判定処理により，左右どちらのノードを実行するのか決定していくことを繰り返し，到達した終端ノードの行動を実行する．判定に使用するデータは，各指標に必要な日数分の始値，高値，安値，終値，出来高の日足データである．

非終端ノードでは，テクニカル指標に基づく判定処理によって左右どちらかのノードを実行する．終端ノードでは，株式の売買行動を実行する．付録 A.1 に終端ノードと非終端ノードを示す．

2.3. 遺伝子操作

戦略木を構成するノードを個体の遺伝子として，交叉や突然変異，淘汰の操作でより高い利益を生み出す個体を作り出していく［7］．

2.3.1. 初期個体生成

表 1 のノードの中から無作為に 1 つを選択する．そのノードが終端ノードであれば，戦略木の生成を終了する．非終端ノードであれば，その非終端ノードの左右にノードを無作為に選択して付け加える．この作業を，全ての非終端ノードの左右に終端ノードが付け加えられるまで繰り返す．初期個体の生成は，初めからノード数の多い木になることを防ぐため，深さ 4 には終端ノードを付け加える．初期個体は N 個体生成する．

2.3.2. 評価

個体の適応度 F を次式のように定義した．

$$F = \frac{\text{総利益}}{\text{総損失}} \times \text{勝率} \times \log(\text{総取引数})$$

ここで，総利益は複数銘柄の複数期間でそれぞれ取引を行った時の利益の合計，総損失は複数銘柄の複数期間でそれぞれ取引を行った時の損失の合計，勝率は総取引数のうち利益を得た取引数の割合である．

複数の期間でそれぞれ取引を行った時の損失が小さく利益が大きい，つまりさまざまな変動で利益を上げることができる個体の適応度が高くなるようになっている．さらに，勝ち取引数をかけることによって勝った取引の多い個体の適応度が高くなるようになっている．

さらに，学習データを銘柄ごとに期間を短く分割することでトレンド等の時系列の特徴を出やすくした状態で学習を行う．期間をより細かく分割し，分割したそれぞれの期間で取引することで，多くの変動パターンを学習できると考えたからである．

2.3.3. ノード削除

戦略木によって行動を決定する際に同じ数字のノードを 2 回目以降実行する場合，そのノードでは左右どちらのノードを実行するのか決まっているため，絶対に到達できないノードが生まれてしまう．例えば，図 1 では，深さ 1 と深さ 2 に 8 番のノードがある．深さ 1 にある 8 番で左のノードを実行する場合，深さ 2 の 8 番も左のノードを実行することになり，その逆側の深さ 3 の 23 番のノード以下には到達できない．これらのノードは売買行動に関係なく適応度に影響を与えないが，ノード数を制限していないためこのようなノードが際限なく増えていく．これでは交叉時，適応度に影響を与える部分が選ばれにくくなり，進化が停滞しやすく効率的な学習が妨げられる．そこで，戦略木がある程度大きくなってから必要ない部分を削除するために，101 世代以降では，毎世代評価で使われなかったノードの削除を行う．ノードの削除を 101 世代以降で行うのは，初期世代から行うと戦略木が大きくならなくなってしまうからである．

2.3.4. 交叉

現世代の全個体から適応度を用いたルーレット選択により親となる 2 個体を選択する．それぞれの個体から無作為に 1 つずつノードを選び，そのノード以下の部分木を交換することで子を 2 個体作り出す．図 2 に交叉の例を示す．これらの操作を繰り返すことで N 個体を作り出す．

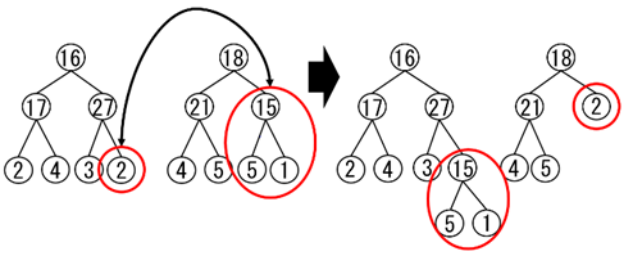


図 2　交叉の例

2.3.5. 突然変異

交叉によってできた N 個体について適応度の順位に基づいて突然変異を施すか決定する．個体 i の突然変異が起きる確率 P_i は次式で定義される．

$$P_i = (Rank_i - 1) \times 0.1 \quad (\%) \qquad (1)$$

ここで，$Rank_i$ は子個体 i の子個体群での適応度の順位である．これにより，順位の高い子個体は突然変異する確率が低くなる．突然変異を施す場合，子個体を構成する各ノードが突然変異率 $p(\%)$ で，終端

ノードなら他の終端ノード，非終端ノードなら他の非終端ノードに無作為に変化させる．

2.3.6. 次世代に残す個体の選択

次世代に残す個体を全 $2N$ 個体から適応度の順位を用いたルーレット選択によって選び出す．こうすることで，個体の多様性を保ちつつ順位の高い個体を次世代に残りやすくする．

2.4. 提案手法の流れ

個体を N 個体用意する．個体はそれぞれ戦略木と現金残高，保有株式数の情報を持ち，1 日 1 回売買行動する．各個体は独立して，戦略木に従い定めた期間，複数銘柄でそれぞれ取引を行う．取引は過去のデータを用いて実市場と同じ値動きを再現した市場で行い，株式の注文は注文を出した日の終値で成立させる．提案手法は以下の流れで行う．

Ⅰ) 初期個体生成
Ⅱ) 初期個体の評価
Ⅲ) 交叉
Ⅳ) 交叉によってできた個体の評価
Ⅴ) 交叉によってできた個体の
　使われないノード削除
Ⅵ) 突然変異
Ⅶ) 突然変異した個体の評価
Ⅷ) 突然変異した個体の使われないノード削除
Ⅸ) 次世代に残す個体の選択

Ⅰ)，Ⅱ) の後 Ⅲ)～Ⅸ)を，決められた世代数まで繰り返す．

3. 結果

本研究で扱う取引銘柄は，東京証券取引所 1 部に上場している 10 業種から，さまざまな変動を学習させるために，出来高が多く長い期間データを取ることができる銘柄を 1 つずつ選び出した，表 2 の 10 銘柄である．

表 2　取引銘柄

業種	銘柄
製造業	キヤノン
水産・農林業	極洋
鉱業	日鉄鉱業
建設業	積水ハウス
金融保険業	野村ホールディングス
不動産業	三井不動産
運輸・情報通信業	KDDI
電気・ガス業	大阪ガス
サービス業	セコム
商業	イオン

個体数 N = 100，世代数 5000，試行回数 50 とし，突然変異率 p は最も良かった p = 0.5 (%)とした．学習期間は 2000 年から 2005 年とした．学習期間の分割の刻みは，2 年，1 年の 2 つである．分割した期間においてそれぞれ元金 10 万円とし，全銘柄で取引を行う．例えば，2 年刻みであれば，1 銘柄につき 6 年間の学習期間を 2000 年 1 月～2001 年 12 月，2002 年 1 月～2003 年 12 月，2004 年 1 月～2005 年 12 月の 3 期間に分け，10 銘柄分で 30 パターンの株価変動で取引し適応度を計算する．テスト期間は 2006 年から 2016 年とし 1 年ごとにテストを行った．

図 3 に学習期間（2000 年 ～ 2005 年）の各世代で適応度が最も高い個体の 1 銘柄あたり 1 年間に換算した利益率の 50 試行平均を示す．学習期間にはさまざまな変動が存在するが，上昇トレンドが多い期間となっている．この図の縦軸は 1 銘柄あたり 1 年間に換算した利益率，横軸は世代数を表し，黒線は既存手法，オレンジ線は 2 年刻みによる提案手法，青線は 1 年刻みによる提案手法の利益率を表す．

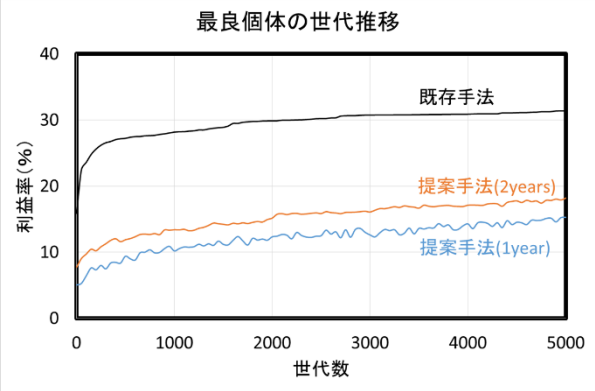


図 3　最良個体の利益率の世代推移（学習期間）

既存手法の戦略木，提案手法の戦略木ともに世代が進むにつれて利益率が上がっており，学習が進んでいることがわかる．学習期間において既存手法の戦略木が最も利益を上げることができている．これは既存手法の戦略木は稼ぎやすい上昇トレンドのみで利益を上げるように進化しており，提案手法の戦略木はさまざまなパターンの変動で利益を上げようと進化しているため上昇しているところで十分に稼ぐことができていないためである．短く分割し学習を行ったものほど利益を上げることができていないのは，より多くのパターンで利益を上げようと進化しており，上昇トレンドで稼ぐことがより難しくなっているためである．

図 4 に 2009 年の各世代で適応度が最も高い個体の 1 銘柄あたり 1 年間に換算した利益率を 50 試行平均した世代推移を示す．2009 年は下降トレンド後，上昇トレンドまたは横ばいの銘柄が多い期間である．縦軸は 1 銘柄あたり 1 年間に換算した利益率，横軸

は世代数を表し，黒線は既存手法，オレンジ線は 2 年刻みによる提案手法，青線は 1 年刻みによる提案手法の利益率を表す．

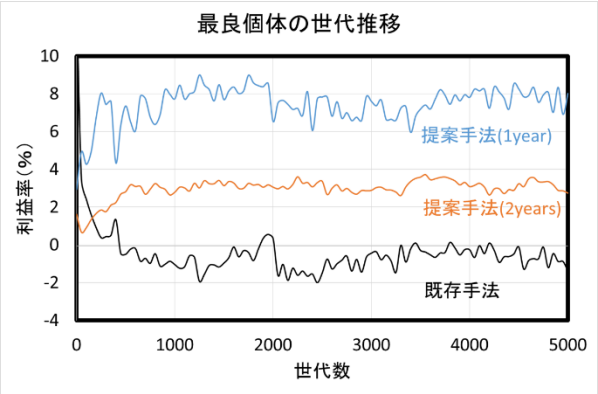


図 4　最良個体の利益率の世代推移（2009 年）

最終世代ではどの刻みによる提案手法も既存手法の利益を上回っており，短く分割し学習を行った 1 年刻みによる提案手法の戦略木が最も利益を上げることができている．提案手法は学習期間では既存手法より利益を上げることができていなかったが，このテスト期間では利益を上げることができている．これは，既存手法は上昇トレンドのみでしか稼ぐことができないように進化したため，テスト期間の変動に対応できず，提案手法はさまざまなパターンの変動で利益を上げようと進化しているためテスト期間の変動にも対応することができたからだと考えられる．

2009 年以外の期間でもテストを行ったところ，提案手法は上昇トレンドでは既存手法に劣るが，他の変動では上回ることができた．詳しい結果と考察については発表時に述べる．

4. 今後の課題

さまざまな変動に対応することができたといったが，上昇トレンドではあまり利益を上げることができておらず，それぞれのトレンドに対応したとは言い難い．現在の木構造ではさまざまなトレンドに対応した戦略の記述は難しいのではないかと考えており，新たな戦略の記述方法の検討を行う必要がある．具体的には，ノードに用いるテクニカル指標によって分岐の数を増やして戦略木を構築するという新たな戦略の記述方法を検討中である．

本研究では，株式を扱ったが為替などの別の金融取引に対しての適用も考えている．株式では出来高のデータを使用することができるが，為替では使用することができない．そのため，出来高の代わりになるデータを用いる必要がある．

参考文献

[1] 松村幸輝，国屋美敬，木村周平: 遺伝的プログラミング手法に基づくエージェントベーストレーダモデル，情報処理学会論文誌，Vol.47，No.9，pp.2869-2886，(2006)
[2] 松村幸輝: 進化計算手法にクラスタリングを応用した株式投資の戦略木最適化モデル，情報処理学会論文誌, Vol.49, No.1, pp.457-475, (2008)
[3] 松村幸輝: Walsh 変換を用いた遺伝的プログラミング手法に基づくファジィ交渉エージェント，情報処理学会論文誌, Vol.46, No.1, pp.267-278, (2005)
[4] 松村幸輝，河元勝 : 行動ファイナンスに基づくデリバティブ取引の GP 最適化モデル，電気学会論文誌 C, Vol.132, No3, pp.455-466, (2011)
[5] 長尾優，森直樹，中島義裕，松本啓之亮 : デイトレードエージェントフレームワークを用いた遺伝的プログラミングによる投資戦略の進化，システム制御情報学会論文誌，Vol.21, No12, pp.400-407, (2008)
[6] 松野弘基，北村信吾，森直樹，松本啓之亮 : 動的なノード変更機構を導入した遺伝的プログラミングによる投資戦略の進化，第 52 回自動制御連合講演会，E3‐4，（2009）
[7] Koza, J : Genetic Programming, MIT Press, (1992)
[8] 吉見俊彦: 決定版チャート分析の真実, 日経ラジオ社, (2003)
[9] 山中康司: 相場で勝つための基本と実践　テクニカル指標の読み方・使い方, (2015)
[10] 藤本壱: 株価チャート分析大全，自由国民社，(2004)

付録

付録 A.1 終端ノードと非終端ノード

	番号	内容
終端ノード	1	資金が許す限り株式を購入，資金がないときは待機
	2	所有している株式をすべて売却，株式がないときは待機
	3	待機
非終端ノード	4	前日の終値が前々日の終値より高ければ左のノードを実行，それ以外であれば右のノードを実行
	5	前日の出来高が前々日の出来高より高ければ左のノードを実行，それ以外であれば右のノードを実行
	6	前々日 STC-slow < STC-fast，前日 STC-slow > STC-fast ならば左のノードを実行，それ以外であれば右のノードを実行
	7	前々日 STC-slow > STC-fast，前日 STC-slow < STC-fast ならば左のノードを実行，それ以外であれば右のノードを実行
	8	%R を用いて 0 が連続して 3 回出た後に前日の値が 20 以上ならば左のノードを実行，それ以外であれば右のノードを実行
	9	%R を用いて 100 が連続して 3 回出た後に前日の値が 80 以上ならば左のノードを実行，それ以外であれば右のノードを実行
	10	RSI > 70% ならば左のノードを実行，それ以外であれば右のノードを実行
	11	RSI < 30% ならば左のノードを実行，それ以外であれば右のノードを実行
	12	短期移動平均と前日の終値の間に+10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	13	短期移動平均と前日の終値の間に-10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	14	中期移動平均と前日の終値の間に+10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	15	中期移動平均と前日の終値の間に-10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	16	長期移動平均と前日の終値の間に+10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	17	長期移動平均と前日の終値の間に-10%以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	18	ゴールデンクロス（短期と中期）であれば左のノードを実行，それ以外であれば右のノードを実行
	19	ゴールデンクロス（中期と長期）であれば左のノードを実行，それ以外であれば右のノードを実行
	20	デッドクロス（短期と中期）であれば左のノードを実行，それ以外であれば右のノードを実行
	21	デッドクロス（中期と長期）であれば左のノードを実行，それ以外であれば右のノードを実行
	22	短期移動平均との間に+2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	23	短期移動平均との間に-2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	24	中期移動平均との間に+2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	25	中期移動平均との間に-2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	26	長期移動平均との間に+2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	27	長期移動平均との間に-2σ以上の乖離があれば左のノードを実行，それ以外であれば右のノードを実行
	28	VR1 ≦ 70% ならば左のノードを実行，それ以外であれば右のノードを実行
	29	VR1 ≧ 450% ならば左のノードを実行，それ以外であれば右のノードを実行
	30	VR2 ≦ 30% ならば左のノードを実行，それ以外であれば右のノードを実行
	31	VR2 ≧ 70% ならば左のノードを実行，それ以外であれば右のノードを実行
	32	WVR ≦ -40% ならば左のノードを実行，それ以外であれば右のノードを実行
	33	WVR ≧ 40% ならば左のノードを実行，それ以外であれば右のノードを実行

使用しているテクニカル指標を以下に示す[８][９][１０].

・STC-slow，STC-fast（ストキャスティクス値）
一定期間の株価の値上がり幅と値下がり幅から株価の転換点を判断するオシレータ系の指標

・RSI
一定期間の株価の値上がり幅と値下がり幅から株価の転換点を判断するオシレータ系の指標

・%R
一定期間の株価の最高値と最安値から株価の転換点を判断するオシレータ系の指標

・VR1，VR2，WVR（ボリュームレシオ）
出来高の推移から株価の転換点を判断する出来高系の指標

人工市場を用いた負の売買手数料（リベート）が株式市場に与える影響の調査

Investigation how negative trading fees (rebates) affect a stock market with an agent-based model

星野　真広[1]*　水田　孝信[2]　八木　勲[1]
Mahiro Hoshino[1]，Takanobu Mizuta[2]，Isao Yagi[1]

[1] 神奈川工科大学情報学部工学科
[1] Department of Information and Computer Sciences, Faculty of Information Technology, Kanagawa Institute of Technology
[2] スパーク・アセット・マネジメント株式会社
[2] SPARX Asset Management Co., Ltd.

Abstract: 現在，米国を中心に市場構成の見直しの一環として，メイカー・テイカー制をめぐる議論が活発化している．メイカー・テイカー制は，市場に注文を供給するメイカーにリベート（負の手数料）を支払い，その注文を消費するテイカーから手数料を取る手数料体系である．メイカー・テイカー制は，リベートによってメイカーから多くの指値注文を得て取引所のシェアが高まると言われているが，市場にどのように影響を与えるのか良くわかっていない．そこで本研究では，リベートの金額を変化させ，メイカー・テイカー制が市場とテイカーの平均購入価格に与える影響を分析し，メイカー・テイカー制の利点と欠点のトレードオフを検証した．その結果，メイカー・テイカー制は効率的な市場を形成するが，その弊害としてテイカーの平均購入価格が上昇することがわかった．

1　はじめに

現在，米国を中心に市場構成の見直しの一環として，メイカー・テイカー制（又はメイカー・テイカー・モデル）をめぐる議論が活発化している [川本 15, 岡田 17]．通常，市場は自市場で取引をする全ての投資家から手数料を徴収することで，利益を上げ運営を行う．しかし，メイカー・テイカー制を採用した市場では約定した注文に対して，市場に注文を提供していた指値注文側にリベート（負の売買手数料）を支払い，その指値注文を消費した成行注文側から手数料を取る．このような売買手数料を取り入れる利点は，多くの指値注文を得て取引のシェアを高める可能性があること，メイカーが調整した価格で指値注文することにより，価格スプレッドが減少し効率的な市場が形成される可能性があることが挙げられる．しかし，実際に市場にどのように影響を与えるのか良くわかっていない．

そこで本研究では人工市場を用いて手数料が変化した際の市場への影響調査を行う．人工市場を用いる理由は，実証研究では実際に利用されていない手数料での取引によって生じる市場への影響調査を行うことは困難であるが，人工市場を利用するとそれら架空の手数料が市場に与える影響についても確認できるからである．

本研究ではメイカー・テイカー制が導入された人工市場を形成し，メイカーが手数料を考慮に入れた取引を行った際の市場への影響を，マーケット・インパクト，ボラティリティ，市場非効率性，出来高の面より分析した．また，テイカーの平均購入価格を求め，効率的な市場の形成とテイカー側の購入金額の増加のトレードオフを検証した．

この際，各投資家成分の重みを学習して売買決定を行うノーマルエージェントが参入する市場に，メイカーとして，自身のポジションを考慮に入れてマーケットメイクを行うポジションマーケットメイカーエージェントを，テイカーとしてアルゴリズムトレードをモデル化したアルゴリズムエージェントを導入し，市場で取引を行わせることで現実の市場に近いモデルを構築した．

本実験の結果，メイカー・テイカー制は効率的な市場を形成するが，その弊害としてテイカーの平均購入価格が上昇することが確認できた．

*神奈川工科大学情報学部情報工学科
〒 243-0292 神奈川県厚木市下荻野 1030
E-mail: s1621085@cce.kanagawa-it.ac.jp

2　人工市場モデル

本研究では益田ら [益田 19] の人工市場モデルを基に，水田ら [水田 14] のアルゴリズムエージェントと草田ら [草田 15] のポジションマーケットメイカーを加えてモデルの構築を行った．元のモデルで使用されているノーマルエージェントはテイカーとメイカーの両方，アルゴリズムエージェントはテイカー，ポジションマーケットメイカーはメイカーの役割を果たす．

本モデルでは，一つの資産のみを取引対象とする．ノーマルエージェントが n 体，アルゴリズムエージェントが l(ただし，$l \leq n$) 体，合計 $n+l$ 体のエージェントがおり，エージェント $j=1$ から，$j=2,3,4,\cdots$ と順番に注文を出す．最後のエージェント $j=n+l$ が注文を出すと，次の時刻にはまたはじめのエージェント $j=1$ が注文を出していく．通常時はノーマルエージェントが注文を行い，ノーマルエージェントが n/l（小数点以下切り捨て）体注文を出すごとに，アルゴリズムエージェントが 1 体注文を出す．ただしアルゴリズムエージェントもノーマルエージェント同様，$k=1,2,3,\cdots,l$ の順に注文を出し，$k=l$ まで到達すると $k=1$ に戻る．また，ポジションマーケットメイカーが 1 体おり，各エージェントが注文を行う前に売りと買いの注文を行う．

時刻 t はノーマルエージェントおよびアルゴリズムエージェント 1 体が注文を出すたびに 1 だけ増える．すなわち，注文をしただけで取引が成立しない場合も時刻 t は 1 ステップ進む．ポジションマーケットメイカーの注文で時刻 t は進まない．

このモデルでの価格決定メカニズムは，買い手と売り手が価格を提示し，両者の提示価格が合致するとその価格での取引が成立するザラバ方式（連続 double auction 方式）とした．

2.1　手数料体系

市場は営利事業であり，その利益は各投資家が取引を行った際の手数料でまかなっている．メイカー・テイカー制を採用した市場の，市場の利益，メイカーの手数料とテイカーの手数料の関係式は式 (1) となる．

$$R_{ex} = C_m + C_t \tag{1}$$

R_{ex} は市場の必要利益，C_m はメイカーの手数料，C_t はテイカーの手数料を表す．ただし，$R_{ex}=0.1\%$ とする．なぜなら，取引所の利益は，市場形成の効率化とテイカーの購入金額増加のトレードオフに本質的な影響を与えないからで，簡潔なモデルを構築するため固定値とした．C_m，C_t 共に正の値ならば手数料を支払うこと，負の値ならば手数料分を受け取ることを意味する．R_{ex}，C_m，C_t は後述するファンダメンタル価格 P_f の割合で示す．

2.2　ノーマルエージェント

ノーマルエージェントは一般投資家を表すエージェントである．各々のエージェントがファンダメンタル価格を参照し投資判断を行うファンダメンタル成分，過去の価格推移を利用して投資判断を行うテクニカル成分，投資家の試行錯誤的な投資判断を表すノイズ成分を保持している．注文価格は，これらの投資戦略成分から算出された期待リターンを基に求められる．また，ファンダメンタル成分，テクニカル成分を学習することで，市場状況に応じて戦略の比重を切り替える．

2.2.1　注文プロセス

ノーマルエージェントは以下の手順に従い，買いと売りの判断を行う．ノーマルエージェント j が時刻 t の時に予想する価格の変化率（予想リターン）${r_e}_j^t$ は式 (2) から求められる．

$${r_e}_j^t = \frac{1}{{w_1}_j^t + {w_2}_j^t + u_j}\left({w_1}_j^t {r_1}_j^t + {w_1}_j^t {r_2}_j^t + u_j \epsilon_j^t\right) \tag{2}$$

ここで，${w_i}_j^t$ は時刻 t におけるノーマルエージェント j の i 項目の重みであり，シミュレーション開始時にそれぞれ 0 から w_{imax} までの一様乱数で決める．右辺のカッコ内の 1 項目の ${w_1}_j^t$ はファンダメンタル価格を参照し投資判断を行うファンダメンタル成分の重み，2 項目の ${w_2}_j^t$ は過去の価格推移を参照し投資判断を行うテクニカル成分の重みである，これら重みは学習プロセスにて変化する．u_j は各エージェントに固有の働きをさせるためのノイズ成分の重みであり，シミュレーション開始時にそれぞれ 0 から u_{max} までの一様乱数で決められ，シミュレーション中は変化しない．これらの重みは互いに独立して変化する．式 (2) の右辺の最初の項にて正規化することでこれら 3 つの重みからくる影響を平準化している．

${r_i}_j^t$ は時刻 t におけるノーマルエージェント j の i 項目の予想リターンであり，買いと売りの判断時に計算される．1 項目の ${r_1}_j^t$ はファンダメンタル成分のリターンでありファンダメンタル価格と 1 期前の市場価格を比較し，市場価格の方が安ければ正，高ければ負の予想リターンを表す．時刻 t におけるノーマルエージェント j のファンダメンタル成分の予想リターン ${r_1}_j^t$ は $\ln\left(P_f/P^{t-1}\right)$ とする．P_f は時間で変化しない一定のファンダメンタル価格である．P^t は時刻 t における市場価格 (取引されなかった時刻では直近取引された価格であり，$t=0$ では $P^t=P_f$ とする)．2 項目の ${r_2}_j^t$

はテクニカル成分の予想リターンであり過去のリターンが正なら正，負なら負の予想リターンを表す．時刻 t におけるノーマルエージェント j のテクニカル成分の予想リターン ${r_2}_j^t$ は $\ln\left(P^{t-1}/P^{t-1-\tau_j}\right)$ とする．τ_j は 1 から τ_{max} までの一様乱数でエージェントごとに決める．ϵ_j^t は時刻 t，エージェント j のノイズ成分であり，平均 0，標準偏差 σ_ϵ の正規分布乱数である．

式 (2) で導いた予想リターンを元に予想価格 ${P_e}_j^t$ を式 (3) で求める．

$$ {P_e}_j^t = P^{t-1} exp\left({r_e}_j^t\right) \tag{3} $$

注文価格 ${P_o}_j^t$ は平均 ${P_e}_j^t$，標準偏差 ${P_\sigma}^t$ の正規分布乱数で決める．${P_\sigma}^t$ は式 (4) で求める．

$$ {P_\sigma}^t = {P_e}_j^t \cdot est \tag{4} $$

est を「ばらつき係数」と呼び，$est\,(0 < est \leq 1)$ の範囲で定める．買いと売りの判断は式 (5) にしたがう．

$$ \begin{aligned} &{P_o}_j^t < {P_e}_j^t \text{なら 1 単位の買い} \\ &{P_o}_j^t > {P_e}_j^t \text{なら 1 単位の売り} \end{aligned} \tag{5} $$

2.2.2 学習プロセス

学習はエージェントごとに注文の直前に行われ，各成分の予想リターン ${r_i}_j^t$（ただし，$i = 1, 2$）の符号の向きと学習期間のリターン ${r_l}^t = \ln\left(P^{t-1}/P^{t-1-t_l}\right)$ の符号の向きとを比較し，${w_i}_j^t$ を式 (6) のように書き換える．

$$ \begin{aligned} &\text{同符号なら，}\quad {w_i}_j^t \leftarrow {w_i}_j^t + k_l \left|{r_l}^t\right| {q_j}^t \left(w_{imax} - {w_i}_j^t\right) \\ &\text{異符号なら，}\quad {w_i}_j^t \leftarrow {w_i}_j^t - k_l \left|{r_l}^t\right| {q_j}^t {w_i}_j^t \end{aligned} \tag{6} $$

ここで，k_l は定数，${q_j}^t$ は時刻 t，エージェント j に与えられる 0 から 1 までの一様乱数である．式 (6) では，価格変化の方向の予測と現実の価格変化の方向が一致した戦略の重みを引き上げ，外れている戦略の重みを引き下げるようにしている．また式 (6) の学習プロセスとは別に，${w_i}_j^t$ を確率 m で 0 から w_{imax} までの一様乱数にて再設定を行う．

2.3 アルゴリズムエージェント

アルゴリズムエージェントはアルゴリズムトレードをモデル化した機関投資家エージェントを想定している．アルゴリズムトレードとは市場の動向に応じて自動的に買い売りや数量の判断を行い，注文を繰り返す取引のことを指す．このエージェントは規則的に注文数 1 の買い注文のみを行う．導入の目的としては平均購入価格を求めるためのマーケット・インパクト (MI) の測定に利用するためである．

アルゴリズムエージェントは発注の際に，注文板から最も安い売り注文（最良売り気配値，なお最も高い買い注文を「最良買い気配値」という）を確認する．最良売り気配値が存在すればその価格からティックサイズ ΔP を加えた価格で買い注文を出す．注文板に最良売り気配値が存在しない場合，注文は行わない．

2.4 ポジションマーケットメイカー

ポジションマーケットメイカーはマーケットメイク戦略をとる機関投資家エージェントを想定している．自身のポジション（保有しているリスク資産数，買いなら正，売りなら負）を考慮に入れ，最良買い気配値と最良売り気配値から注文基準価格を求め，期待する 1 取引あたりの利益率である提示スプレッド θ_{pm} を，売りなら加えた価格，買いなら減じた価格で取引を行う．草田ら [草田 15] は市場は 2 市場を考慮してモデル化していたが，本研究では 1 市場を対象としてモデル化した．

時刻 t における取引市場の最良売り気配値 $P^{t,sell}$，最良買い気配値 $P^{t,buy}$，ポジションマーケットメイカーの提示スプレッドを θ_{pm}，時刻 t と $t+1$ の間にポジションマーケットメイカーが抱えるポジションを s_{pm}^t，ポジション考慮度を w_{pm} とすると，買い注文価格 $P_{o,pm}^{t,buy}$ と売り注文価格 $P_{o,pm}^{t,sell}$ は式 (7)，式 (8)，式 (9) で決定される．

$$ P_{o,pm}^{t,buy} = P_{fv,pm}^t - \frac{1}{2}\left(P_f \cdot \theta_{pm}\right) \tag{7} $$

$$ P_{o,pm}^{t,sell} = P_{fv,pm}^t + \frac{1}{2}\left(P_f \cdot \theta_{pm}\right) \tag{8} $$

$$ P_{fv,pm}^t = \left(1 - w_{pm}\left(s_{pm}^t\right)^3\right) \cdot \frac{1}{2}\left(P^{t,buy} + P^{t,sell}\right) \tag{9} $$

また，ポジションマーケットメイカーの成行注文を防ぐために価格を一定の値に抑える．制約を式 (10) に示す．

$$ \begin{aligned} P_{o,pm}^{t,buy} &\geq P^{t,sell} \\ P_{o,pm}^{t,sell} &\leq P^{t,buy} \end{aligned} \tag{10} $$

これらの制約時のポジションマーケットメイカーの発注価格は式 (11)，式 (12) のようになる．これにより買い注文と売り注文の価格が逆転することも防ぐことができる．

$$ P_{o,pm}^{t,buy} = \begin{cases} P^{t,sell} - \Delta P & \left(P_{o,pm}^{t,buy} \geq P^{t,sell}\right) \\ P_{o,pm}^{t,sell} - \left(P_f \cdot \theta_{pm}\right) & \left(P_{o,pm}^{t,sell} \leq P^{t,buy}\right) \end{cases} \tag{11} $$

$$P_{o,pm}^{t,sell} = \begin{cases} P_{o,pm}^{t,buy} + (P_f \cdot \theta_{pm}) & \left(P_{o,pm}^{t,buy} \geq P^{t,sell}\right) \\ P^{t,buy} + \Delta P & \left(P_{o,pm}^{t,sell} \leq P^{t,buy}\right) \end{cases} \quad (12)$$

本実験ではポジションマーケットメイカーに，1 取引（1 単位の売り注文と買い注文）で生じる期待利益を設ける．ポジションマーケットメイカーは期待利益とメイカーの手数料 C_m を考慮に入れ提示スプレッド θ_{pm} を調節していく．θ_{pm} の求め方を式 (13) に示す．

$$\theta_{pm} = R_m + 2C_m \quad (13)$$

R_m はポジションマーケットメイカーが 1 取引あたりに必要な期待利益である．モデルを簡潔にするため，本実験では $R_m = 0.3\%$ の固定値とした．本研究では，R_m をファンダメンタル価格 P_f の割合で示す．C_m に係数 2 がついていることに注意する．これは 1 取引で，売り注文が約定した手数料と，買い注文が約定した手数料を支払うことを意味している．以上の条件の下で，ポジションマーケットメイカーの手数料を変化させたときの提示スプレッドとテイカーの手数料の変化の様子を表 1 に示す．

表 1: メイカーの手数料が変化したときの提示スプレッドとテイカーの手数料（ただし，$R_m = 0.300\%, R_{ex} = 0.100\%$）

C_m	θ_{pm}	C_t
0.100%	0.500%	0.000%
0.075%	0.450%	0.025%
0.050%	0.400%	0.050%
0.025%	0.350%	0.075%
0.000%	0.300%	0.100%
-0.025%	0.250%	0.125%
-0.050%	0.200%	0.150%
-0.075%	0.150%	0.175%
-0.100%	0.100%	0.200%
-0.125%	0.050%	0.225%
-0.140%	0.020%	0.240%
-0.145%	0.010%	0.245%

3 実験

本実験ではメイカーが手数料を考慮して取引を行う．表 1 のようにメイカーの手数料 C_m を-0.145%，-0.140%，そして-0.125%から 0.100%の 0.025%刻みと変化させたときの，マーケット・インパクト，ボラティリティ，市場非効率性，出来高を取得した．本論文では，手数料をファンダメンタル価格に対する割合で表す．例えば，ファンダメンタル価格が 10,000，メイカーの手数料が-0.100%である時，メイカーが支払う手数料は-10.0 となる．

各パラメータの値は $n = 990$，$l = 10$，$w_{1max} = 1$，$w_{2max} = 10$，$u_{max} = 1$，$\tau_{max} = 10,000$，$\sigma_\epsilon = 0.06$，$est = 0.003$，$\Delta P = 1.0$，$P_f = 10,000.0$，$t_l = 10,000$，$t_c = 20,000$，$k_l = 4.0$，$m = 0.010$，$w_{pm} = 0.00000005$ とする[1]．

3.1 マーケット・インパクト

アルゴリズムエージェントの平均購入価格を求めるためにマーケット・インパクト (MI) を利用する．MI は自分自身の注文によっていかに市場価格に影響を与えたかを示す．今回は，アルゴリズムエージェントがファンダメンタル価格よりどれだけ高い価格で取引を行ったかを MI と定義した．MI を式 (14) に示す．

$$MI = \frac{1}{n_{buy}} \sum_{j=1}^{n_{buy}} \frac{p_{buy}^j - P_f}{P_f} \quad (14)$$

この MI は実証研究では観測できないファンダメンタル価格 P_f を使用しており，人工市場でしか用いることのできない定義である．ここで，n_{buy} はシミュレート全期間を通してアルゴリズムエージェントが買った数量，p_{buy}^j は各々の注文が成立した価格である．アルゴリズムエージェントが存在しない場合，取引価格の平均はファンダメンタル価格とほぼ同じになることが明らかになっている [水田 14]．そのため，MI を測定して値が高くなればアルゴリズムエージェントの取引が市場に影響を及ぼしているといえる．$MI = 0$ となれば，アルゴリズムエージェントの取引は影響を及ぼしていないことになる．

MI は 1 度に大量の注文を出すことや，注文板上にばらつきがあると値が大きくなりやすい．本モデルでは 1 度に大量の注文を出すことはないが，アルゴリズムエージェント，ノーマルエージェント，特にポジションマーケットメイカー取引により注文板の状態は大きく変化する．これにより MI の測定で取引の影響を分析することができる．

MI はファンダメンタル価格からの割合で表す．例えば $MI = 0.15\%$，$P_f = 10,000$ の場合．アルゴリズムエージェントはファンダメンタル価格から 15 離れた 10,015 で取引を行っている．

[1] モデルの妥当性は実証研究 [Sewell 06, Cont 01] で得られている統計的性質 (Stylized fact) が満たされているかで判断した

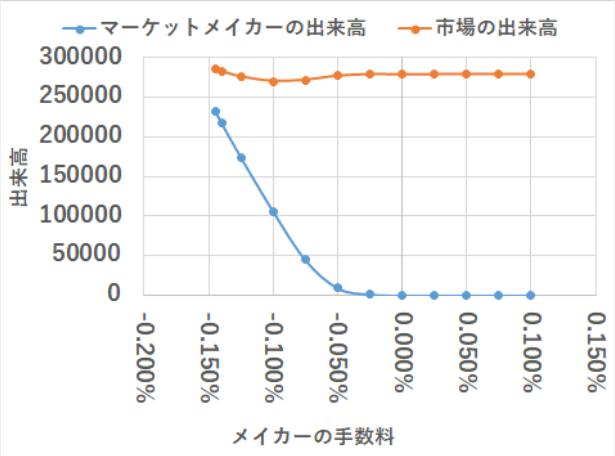


図 1: メイカーの手数料と各々の出来高

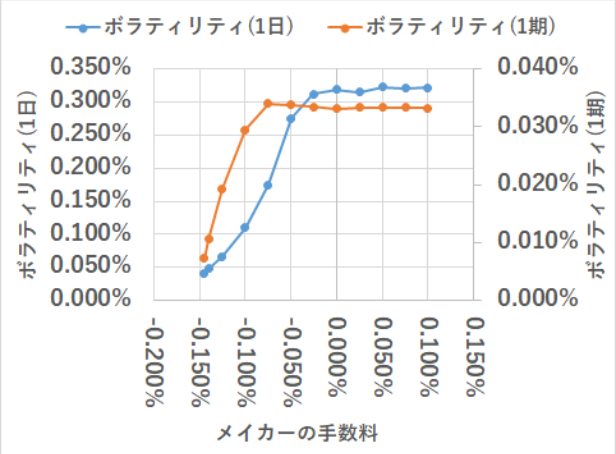


図 2: メイカーの手数料とボラティリティ

3.2 市場非効率性

市場の効率性を測定する指標として水田ら [水田 15] の市場非効率性 M_{ie} を用いる．式 (15) に市場非効率性の求め方を示す．

$$M_{ie} = \frac{1}{t_e} \sum_{t=0}^{t_e} \frac{|P^t - P_f|}{P_f} \tag{15}$$

この市場非効率性も MI と同様に実証研究では観測できないファンダメンタル価格 P_f を使用しており，人工市場でしか用いることのできない定義である．現実市場とは異なりファンダメンタル価格 P_f が明確なため，推定ではない理想的な市場の効率性を測定できる．M_{ie} は 0 以上の値をとり，0 ならば完全に効率的，値が大きくなるにつれて非効率的であることを表している．市場非効率性はファンダメンタル価格 P_f に対する市場価格の平均乖離度で表している．

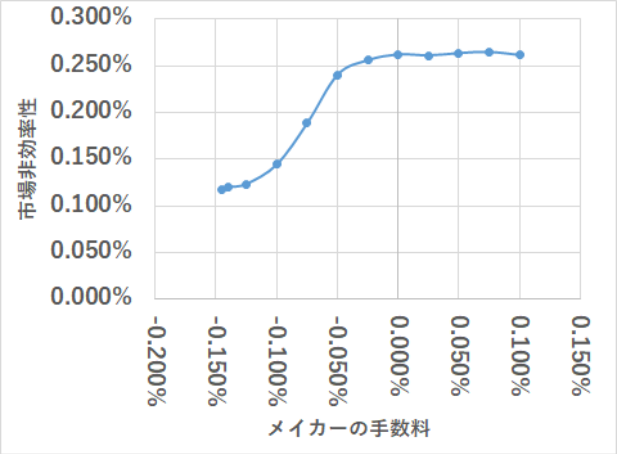


図 3: メイカーの手数料と市場非効率性

4 結果と考察

実験の結果，メイカーの手数料が低下がするにつれて，ポジションマーケットメイカーの出来高は増加，ボラティリティは低下，市場非効率性は低下，マーケット・インパクトは減少した．

4.1 ポジションマーケットメイカーの出来高

メイカーの手数料が低下すると，ポジションマーケットメイカーの出来高は増加した（図 1 参照）．メイカーの手数料が低い環境，つまりメイカーがリベートを受け取りやすくなる環境下では，ポジションマーケットメイカーは期待利益をリベートで補えるため，提示スプレッドを狭めることができる．提示スプレッドが狭まると注文価格が市場価格に近付き，約定しやすい価格になるため，ポジションマーケットメイカーの出来高は増加する．

メイカーの手数料が低下し負になる，すなわちメイカーがリベートを受け取れる環境で出来高を増やすことができれば，その分リベートを受け取る機会が増える．すなわち，メイカーの利益につながる．よって，市場がリベートを支払うことはメイカーの利益につながる．

4.2 ボラティリティ

ボラティリティは 1 日毎（1 日 = 20000 期とする）と，1 期毎のものを取得した．メイカーの手数料が低下すると，ボラティリティは低下する結果を得られた（図 2 参照）．メイカーがリベートを受け取れる環境では，メイカーの期待利益はリベートで補えるため，ポジションマーケットメイカーは狭い提示スプレッドを提供できる．ポジションマーケットメイカーが提示ス

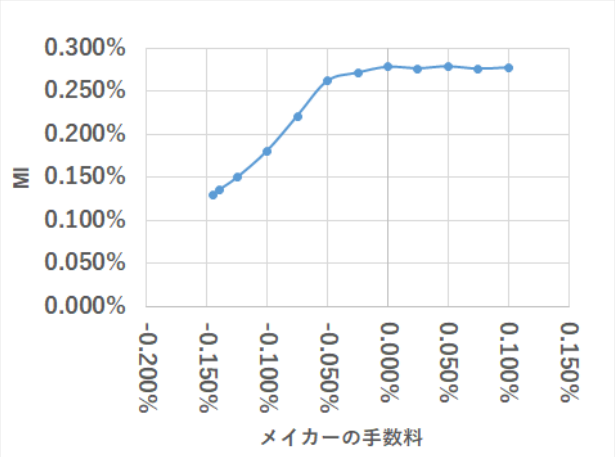


図 4: メイカーの手数料とマーケット・インパクト

プレッドを狭めると，ノーマルエージェントはポジションマーケットメイカーの注文と約定しやすくなる．ポジションマーケットメイカーの注文と約定しやすくなることで市場価格がメイカーの狭い提示スプレッドの間で決まるようになり，価格変動が抑えられる．よって，ボラティリティが低下した．

4.3 市場非効率性

メイカーの手数料が低下すると，市場非効率性は低下した（図 3 参照）．メイカーがリベートが受け取れる環境の場合，市場価格がファンダメンタル価格から離れるような値動きをしたとしても，ポジションマーケットメイカーの形成する狭い提示スプレッド付近で注文が約定される．これにより市場価格の変動幅は小さいものになる．市場価格の変動幅が小さければ，ノーマルエージェントが市場価格がファンダメンタル価格より離れていることを学習し，注文価格を矯正するまでの期間中の市場非効率性への影響は小さくすむ．よって，全体の市場非効率性は減少する．

4.4 マーケット・インパクト

メイカーの手数料が低下すると，マーケット・インパクトは減少した（図 4 参照）．リベートが受け取れない（メイカーの手数料が正の）環境では，ポジションマーケットメイカーが求める期待利益を得るためには広い提示スプレッドを提供する必要が生じる．提示スプレッドが一定を超えるとポジションマーケットメイカーの注文は，ノーマルエージェントの注文よりも買いなら安い，売りなら高い価格になる．すると，アルゴリズムエージェントの注文はノーマルエージェントの指値注文と約定しやすくなる．ノーマルエージェントの指値注文がまばらだと，アルゴリズムエージェントの成行注文によって市場価格が大きく動いてしまうことがある．

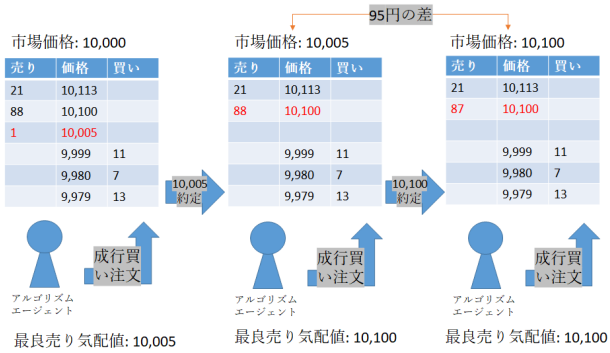


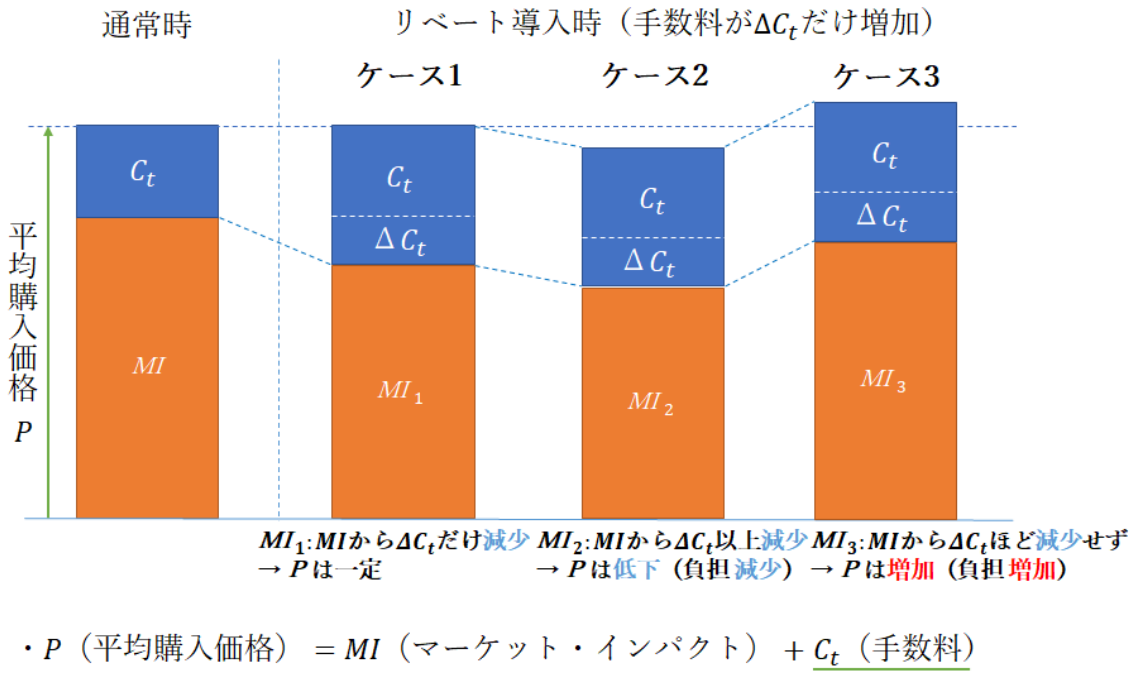
図 5: アルゴリズムエージェントによる価格上昇

図 5 を用いて具体例を説明する．注文板上の売り注文として最良売り気配値が 10,005，次の注文が 10,100 であるとき，まずアルゴリズムエージェントが成行買い注文を出すと 10,005 の注文が約定し市場価格が 10,005，最良売り気配値が 10,100 になる．次にアルゴリズムエージェントが取引するときまでにこの 10,005 付近に売り注文が入れば影響は少ないが，その前に買い注文が来て約定してしまうと市場価格が 10,100 となり市場価格が 95 も動いてしまう．このように，ポジションマーケットメイカーが機能しないとアルゴリズムエージェントは自身の注文の影響で取引価格が上昇させてしまうことがある．

リベートが受け取れる環境では，ポジションマーケットメイカーの注文はアルゴリズムエージェントの成行買い注文と約定しやすくなる．ポジションマーケットメイカーの注文と約定することで，市場価格がポジションマーケットメイカーの注文の提示スプレッドの間で安定し，大きな価格変動は生じにくくなる．

4.5 テイカーの平均購入価格

図 6 はテイカーの平均購入価格とマーケットインパクト，テイカーの手数料（=メイカーのリベート）の関係を表したものである．テイカーの平均購入価格は，テイカーの手数料の増加分よりマーケット・インパクトが減少していると，通常時よりも低くなる（図 6 のケース 2 参照）．つまり，アルゴリズムエージェントは，マーケットメイカーのマーケットメイクによってテイカーの手数料の増加分よりマーケット・インパクトを抑えられれば有利になる．一方，テイカーの平均購入価格は，テイカーの手数料増加分よりマーケット・インパクトが増加していると通常時よりもは高くなる（図 6 のケース 3 参照）．



・P（平均購入価格）＝MI（マーケット・インパクト）＋C_t（手数料）

図 6: テイカーの平均購入価格とマーケットインパクト，テイカーの手数料の関係の関係

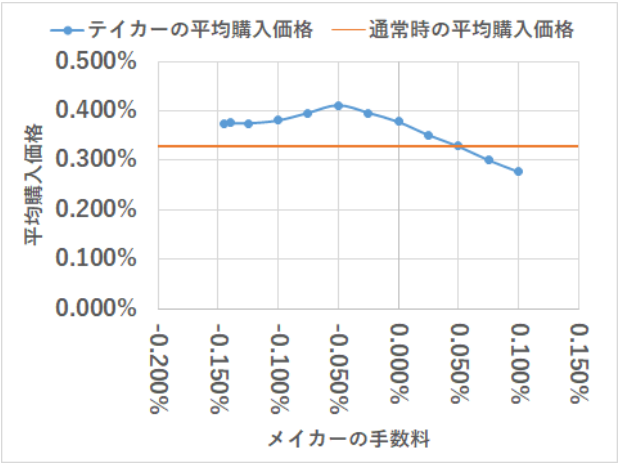


図 7: テイカーの平均購入価格

図 7 から，通常時（メイカー・テイカー制非導入時）のテイカーの平均購入価格は 0.327%で，リベート導入後（メイカーの手数料が負）の平均購入価格は通常時の平均購入価格より全体的に高くなっていることが分かる．よってアルゴリズムエージェントにとって不利な状況になっている．

以上の議論より次の知見が得られた．メイカー・テイカー制は取引のしやすい市場を提供する．なぜなら，リベートが上昇につれて価格変動が抑制されるため市場参加者が取引しやすくなるからである．しかし，その弊害としてテイカーの平均購入価格が増加してしまう．よって，リベートの導入は，市場とメイカーには有利になるが，テイカーには不利になると思われる．

5　まとめ

本研究では人工市場を用いてメイカー・テイカー制が導入された人工市場を形成し，メイカーが手数料を考慮に入れた取引を行った際の市場への影響をマーケット・インパクト，ボラティリティ，市場非効率性，出来高の面より分析した．また，テイカーの平均購入価格を求め，効率的な市場の形成とテイカー側の購入金額の増加のトレードオフを検証した．

その結果，リベートの導入はマーケット・インパクトの減少，市場の効率化，出来高の上昇といった効果を上げうることが確認できた．しかし，マーケット・インパクトの減少分よりもテイカーの手数料が増え，結果としてアルゴリズムエージェントは高い購入価格で購入させられてしまっていることが実験結果より明らかになった．テイカーの手数料は成行注文を出すすべての投資家に課せられ，アルゴリズムエージェント（機関投資家）のほかにノーマルエージェント（個人投資家）も支払うことになる．このことにより，リベートの導入は市場とメイカーには有利になるが，テイカー（他の投資家）たちは不利になりうることが判明した．

今後の課題として以下が挙げられる．今回はマーケット・インパクト，ボラティリティ，市場非効率性，出来高の面で市場の分析を行った．しかしながら，市場参加者が実際に取引をしやすくなったかといった議論は行われていない．そのため市場の取引のしやすさとして用いられる市場流動性の指標 (Volume，Tightness，Resiliency，Depth) を用いて，メイカー・テイカー制を導入すると流動性にどのような変化がみられるか確認していく．

参考文献

[Cont 01] Cont, R.: Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance*, Vol. 1, pp. 223–236 (2001)

[Sewell 06] Sewell, M.: Characterization of financial time series (2006)

[益田 19] 益田裕司, 水田孝信, 八木勲 : 人工市場を用いた金融市場流動性に影響を与える要因の調査, JPX ワーキング・ペーパー, Vol. 29, (2019)

[岡田 17] 岡田功太, 齋藤芳充:米国株式市場のメイカー・テイカー・モデルを巡る議論-流動性向上策としてのリベートの功罪-, 野村資本市場クォータリー (2017)

[水田 14] 水田孝信, 小杉信太郎, 楠本拓矢, 松本渉, 和泉潔 : ダーク・プールは金融市場を安定化しマーケット・インパクトを低減させるか? 人工市場シミュレーションを用いた検証 , 第 12 回金融情報学研究会 (2014)

[水田 15] 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 (2015)

[川本 15] 川本隆雄 : 米国における株式市場構造改革議論とその行方, 月刊資本市場, Vol. 11, No. 363 (2015)

[草田 15] 草田裕紀, 水田孝信, 早川聡, 和泉潔 : 保有資産を考慮したマーケットメイク戦略が取引所間競争に与える影響: 人工市場アプローチによる分析, 人工知能学会論文誌, Vol. 30, No. 5, pp. 675–682 (2015)

裁定取引へのリベートは流動性を向上させるか？ —人工市場によるシミュレーション分析—

Do Rebates to Arbitrage Trades Improve Liquidity?
— Investigation by Agent-Based Model —

水田孝信 *1
Takanobu Mizuta

*1 スパークス・アセット・マネジメント株式会社
SPARX Asset Management Co., Ltd.

通常，株式の売買を行った投資家は取引所に手数料を支払う．しかしながら近年米国を中心に，対当する注文を待っていた注文（指値注文，メイカー）側にリベート (負の売買手数料) を支払い，対当させた注文（成行注文，テイカー）側から手数料を取る，“メイカー・テイカー制” とよばれる手数料体系をとる取引所が増えている．流動性を供給するメイカー側にリベートを差し出すことにより，流動性を供給するインセンティブを与え，より多くの流動性を供給してもらうのが目的である．しかし，その効果やメカニズムはどのようなものかは分かっていない．そこで本研究では，人工市場（金融市場のエージェントベースドモデル）を用いて，リベートも含め，裁定取引にかかるコストが流動性に与える影響を分析した．その結果，ボラティリティよりコストが小さければ裁定の機会が訪れやすく，裁定エージェントの売買が増え，ETF と株式の価格の乖離が小さくなることが分かった．また，コストがマイナスになっても大きな変化が見られなかった．このことは，手数料をとられるかリベートを受け取れるかの違いは，コストがボラティリティに比べ高いか低いかに比べ，重要ではないことを示している．さらに，売買数量にかかわらず注文数量を一定にした場合は，株式の売買量と板の厚さは負の相関となり実証研究とあわなかった．一方で，売買数量の増加に応じて注文数量を増やした場合は，株式の売買量と板の厚さは正の相関となり実証研究と整合的な結果が得られた．このことは，売買の量が増えればより多くの量を注文する投資家の存在が，流動性の統計的性質を決める重要なメカニズムである可能性を示している．

1. はじめに

通常，株式の売買を行った投資家は取引所に手数料を支払う．しかしながら近年米国を中心に，対当する注文を待っていた注文（指値注文，メイカー）側にリベート (負の売買手数料) を支払い，対当させた注文（成行注文，テイカー）側から手数料を取る，“メイカー・テイカー制” とよばれる手数料体系をとる取引所が増えている [川本 15, 岡田 17]．流動性を供給するメイカー側にリベートを差し出すことにより，流動性を供給するインセンティブを与え，より多くの流動性を供給してもらうのが目的である *1．しかし，その効果やメカニズムはどのようなものかは分かっていない．

実証研究のみではこのようなまだ導入したことがない手数料体系を調べたり，その変更の純粋な効果やメカニズムを議論したりするのは困難である．このような議論をするのにすぐれた手法として，コンピュータ上で仮想的にその状況を作り出し検証する，人工市場（金融市場のエージェントベースドモデル）を用いたシミュレーションがある *2．これまでの伝統的な経済学で使われてきた手法にはない強みがあるとして，Nature と Science に人工市場モデルに期待を寄せる論考が掲載されている [Farmer 09, Battiston 16]．[Bookstaber 17] でも，エージェントベースドモデルを使えば，これまでの伝統的な経済学では十分分析できなかった金融危機を分析できる可能性があると主張している．

そして，人工市場モデルを用いたシミュレーション研究はバブルや金融危機の発生メカニズムの解明に貢献したことはもちろん，現実の金融市場の規制や制度，ルールの変更の議論に多くの貢献をした [水田 19b, Mizuta 19c]*3．また，JPX ワーキングペーパー *4 においても，人工市場モデルを用いて制度変更を検討した研究を数多く公表している．

[水田 19a] は 2 つの株式とそれら合計と同じ価値のある 1 つの ETF(Exchange Traded Funds, 上場投資信託) という 3 つの証券（図 1）とこれらの証券間の裁定取引を行うエージェントを実装した人工市場モデルを用いて，株式と ETF の裁定取引にかかる手数料などのコストによって流動性がどのように変化するか調べた．しかし，リベートは調査していない．さらに裁定取引の対象となる株式の流動性は，売買が増えるにつれて待機注文の量（板の厚さ）が減少しており，実証分析 [Muranaga 99, Chung 05] とあわない結果を出している．

そこで本研究では，[水田 19a] の人工市場モデルに，売買数量を調整できる仕組みを実装して実証分析と整合的な結果を出力するモデルを構築し，[水田 19a] では調べられていなかったリベートも含め，裁定取引にかかるコストが流動性に与える影響を分析する．

2. 人工市場モデル

[Chiarella 02] では，シンプルでありながら，実証分析で得られた長期間に存在する価格変動の統計的性質 (stylized fact) を再現できるエージェントモデルの構築に成功している．[水田 13] では，[Chiarella 02] のモデルをベースにモデルを構築し，[Chiarella 02] のモデルでは再現されていなかったミリ秒からマイクロ秒といった短い時間スケールでの性質 (マーケット・マイクロ・ストラクチャー) も再現した．[水田 19a] は [水田 13] の人工市場モデルをベースに，2 つの株式とそれら合計と同じ価値のある 1 つの ETF という 3 つの証券に拡張（図

連絡先: mizutata@gmail.com
https://mizutatakanobu.com
当日の発表スライドは以下にある．
https://mizutatakanobu.com/2020031415.pdf

*1 株式などの流動性は社会の発展に非常に重要である．詳しい解説として [水田 18] がある．

*2 優れたレビューとして，[LeBaron 06, Chen 12, Todd 16, 和泉 17b] がある．

*3 その他にもレビューとして，[水田 14, Mizuta 16, 和泉 17a] がある．

*4 https://www.jpx.co.jp/corporate/research-study/working-paper/index.html

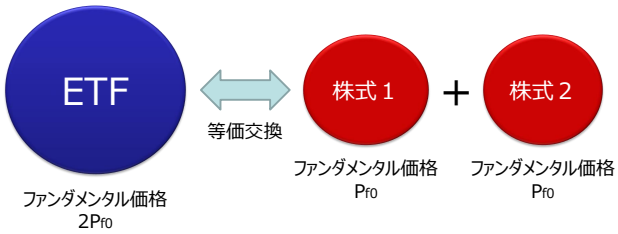


図 1: ETF1 株は，株式 1 の 1 株と株式 2 の 1 株の計 2 株と交換できる

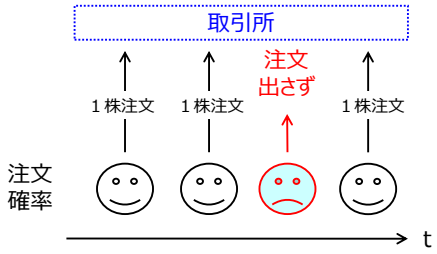


図 2: 注文は確率 $k_h(0 < k_h < 1)$ で実行

1）し，これらの証券間の裁定取引を行うエージェントを実装したモデルを構築した．

本研究では，[水田 19a] の人工市場モデルに，売買数量を調整できる仕組みを実装して実証分析と整合的な結果を出力するモデルを構築し，[水田 19a] では調べられていなかったリベートも含め，裁定取引にかかるコストが流動性に与える影響を分析する．本研究の目的にはモデルがシンプルであることはとても重要である．本モデルの構築の基本理念は [水田 19b, Mizuta 19c] や付録 “モデル構築の基本理念” 参照．

本モデルは 2 つの株式とそれら合計と同じ価値のある 1 つの ETF という 3 つの証券 $h(h = 0$ が ETF, $h = 1$ が株式 1, $h = 2$ が株式 2) を取引対象とする（図 1）．いずれの証券 h も価格決定メカニズムは，ザラバ方式（continuous double auction）[*5] とし，独立して価格の決定・売買が行われる．注文できる価格の変化幅の最小単位 (呼値の刻み) は δP とし，注文価格のそれより小さい端数は，買い注文の場合は切り捨て，売り注文の場合は切り上げる．

いずれかの証券 h のみを売買する多数（各証券に n 体，全部で $3n$ 体）のノーマルエージェントと，3 つの証券に対して裁定取引を行う 1 体の裁定エージェントが存在する．いずれのエージェントも保有する証券の数量に制限はなく (キャッシュが無限大)，マイナスの保有数量 (空売り) にも制限はない．

2.1 ノーマルエージェント

各証券 h に対してそれぞれ n 体のノーマルエージェントが売買に参加する．ノーマルエージェントは，実際の市場の価格形成の性質を再現するために導入するものであり，stylized fact や高頻度取引にかかわる統計量を再現するなるべくシンプルな，ごく一般的な投資家をモデル化したものとした．エージェント番号 $j = 1$ から順番に $j = 2, 3, 4, \ldots$ と注文を出す．$j = n$ が注文を出すと，次の時刻にはまた $j = 1$ から注文を出し繰り返される．注文数量は常に 1 株と一定とする．

また，各証券 $h(h = 0$ が ETF, $h = 1$ が株式 1, $h = 2$ が株式 2) の注文の量（流動性）を調整するため，注文は $k_h(0 < k_h < 1)$ の割合でしか実行されず，実行されなかったエージェントはなにもせず次のエージェントに順番が移る．

注文価格 $P^t_{o,j}$，売り買いの別を以下のように決める．各証券の時刻 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，

$$r^t_{e,j} = (w_{1,j} \log \frac{P_f}{P^t} + w_{2,j} \log \frac{P^t}{P^{t-\tau_j}} + w_{3,j}\epsilon^t_j)/\Sigma^3_i w_{i,j} \quad (1)$$

とする．ここで，$w_{i,j}$ はエージェント j の i 項目の重みであり，シミュレーション開始時に，それぞれ 0 から $w_{i,max}$ まで一様乱数で決める．log は自然対数である．P_f は時間によらず一定のファンダメンタル価格であり株式 1 または 2 の場合 $P_f = P_{f0}$, ETF の場合 $P_f = 2P_{f0}$ である．P^t は各証券それぞれの取引価格，ϵ^t_j は各証券の時刻 t，エージェント j の乱数項であり，平均 0，標準偏差 σ_ϵ の正規分布乱数である．τ_j はシミュレーション開始時に 1 から τ_{max} までの一様乱数でエージェントごとに決める [*6]．

式 (1) の第 1 項目はファンダメンタル価格と比較して安ければプラスの予想リターンを高ければマイナスの予想リターンを示す，ファンダメンタル価値を参照して投資判断を行うファンダメンタル投資家の成分である．第 2 項目は過去のリターンがプラス (マイナス) ならプラス (マイナス) の予想リターンを示す，過去の価格推移を参照して投資判断を行うテクニカル投資家の成分であり，第 3 項目はノイズを表している．

予想リターン $r^t_{e,j}$ より予想価格 $P^t_{e,j}$ は，

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

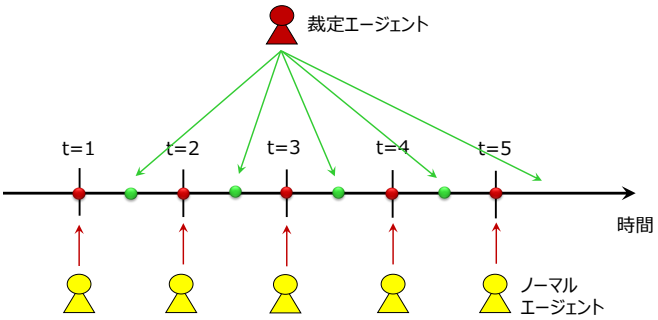


図 3: 裁定エージェントはいつでも注文を出したり，変更したりできるとする

	ETF			株式 1			株式 2	
売り	価格	買い	売り	価格	買い	売り	価格	買い
7	20300		30	10400		50	10400	
10	20200		44	10300		70	10300	
	20100		70	10200		90	10200	
	20000		134	10100		116	10100	
→	19900	1	→	10000	120	→	10000	154
	19800	10		9900	88		9900	60
	19700	6		9800	52		9800	55
	19600	4		9700	25		9700	31

図 4: 裁定取引の例

*5 ザラバ方式は，売り手と買い手の双方が価格を提示し，売り手と買い手の提示価格が合致するとその価格で直ちに取引が成立する方式である [東証 15].

*6 ただし，$t < \tau_j$ のときは第 2 項目を 0 とした．

表 1: Case A, Case B でも用いる株式 1・2 の注文比率 k_1, k_2

コスト	Case A での k_1, k_2 (k_A)	Case A での 20000 期あたり平均売買数量 V_c	V_c/V_0	Case B で使う k_1, k_2 ($k_{B,c} = k_A \times V_c/V_0$)
-0.025%	50%	2955	1.0330	51.65%
-0.010%	50%	2953	1.0323	51.61%
-0.005%	50%	2951	1.0317	51.58%
0.000%	50%	2949	1.0308	51.54%
0.005%	50%	2947	1.0301	51.50%
0.010%	50%	2944	1.0291	51.46%
0.025%	50%	2936	1.0264	51.32%
0.050%	50%	2922	1.0217	51.08%
0.100%	50%	2901	1.0141	50.71%
0.250%	50%	2876	1.0054	50.27%
0.500%	50%	2863	1.0010	50.05%
裁定取引なし	50%	2860 (V_0)	1.0000	50.00%

で求まる．注文価格 $P^t_{o,j}$ は平均 $P^t_{e,j}$，標準偏差 P_σ の正規分布乱数で決める．ここで，P_σ は定数である．そして，売り買いの別は予想価格 $P^t_{e,j}$ と注文価格 $P^t_{o,j}$ の大小関係で決める．すなわち，

$$\begin{array}{l} P^t_{e,j} > P^t_{o,j} \text{なら1株の買い} \\ P^t_{e,j} < P^t_{o,j} \text{なら1株の売り} \end{array} \quad (3)$$

とする [*7]．注文を行ってからキャンセル時間 t_c だけ経過した注文はキャンセルする．

2.2 裁定エージェント

ETF1 株は，株式 1 の 1 株と株式 2 の 1 株の計 2 株と等価であり，各証券の取引価格がいくらであったとしても交換できる．そのため，ETF を買った価格が株式 1 と 2 を売った価格の合計より小さければ，その差額が利益となる（その逆もしかりである）．というのも，安く買った ETF を株式 1 と 2 に交換して高く売ればその差額が利益となるからである．このような価格差を利益とする取引は裁定取引とよばれており，裁定エージェントは裁定取引を行う．裁定エージェントはいつでも注文を出したり，変更したりできるとする (図 3)．

図 4 は，各証券の注文状況の例である．株式 1 と 2 の最も高い買い注文の価格の合計は 20000(=10000+100000) である．一方，ETF の最も高い買い注文の価格は 19800 であり，19900 には買い注文は入っていなかったとする．裁定エージェントは ETF へ 19900 の 1 株の買い注文（赤字）を入れ待機する．もし，この注文が成立し，ETF1 株を手に入れたら，この ETF1 株を株式 1，2 それぞれ 1 株ずつと交換し，株式 1 と 2 を各 10000 で 1 株づつ売れば，差額の 100 が利益となる．もちろん逆の裁定取引も可能である．すなわち，まず ETF を借りてきて高く売り，株式を安く買って ETF に交換して返し，その差額が利益となる．

以上の例では，取引にかかるコストをゼロとしたが，実際にはコストがかかる．本モデルでは上記の一連の取引にかかるすべての手数料などのコストと，1 回の裁定取引で必要な利益額の合計を $C = c \times P_{f0}$ と定義する．C には必要な利益も含んでいるので，上記の価格差が C となる取引ができる場合，裁定取引を行うことができる．また，流動性が低い ETF の取引

*7 ただし，$t < t_c$ のときは十分な板の厚さを確保するため，$P_f > P^t_{o,j}$ なら 1 株の買い，$P_f < P^t_{o,j}$ なら 1 株の売りとする．また，$P^t = P_f$ とする

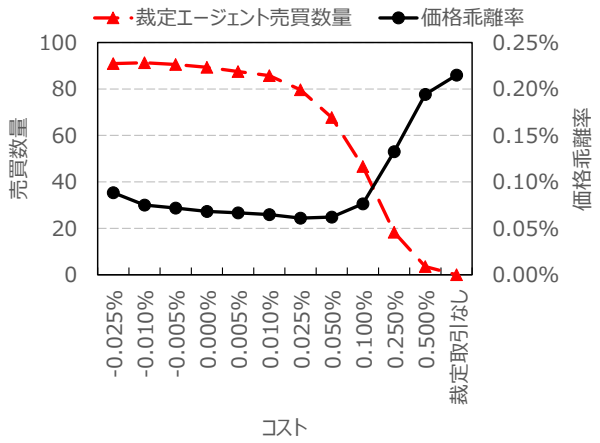


図 5: 価格乖離率と裁定エージェントの売買株数 (Case A : $k_1 = k_2 = 50\%$ で固定)

で特別に大きなリベートが得られる場合も想定し，C は負の値もとることができるとする．

今，株式 1，2，および ETF の最も高い買い注文の価格をそれぞれ B_1, B_2, B_{ETF}，それらの最も安い売り注文の価格をそれぞれ S_1, S_2, S_{ETF} とする．裁定エージェントは，$B_{ETF} < B_1 + B_2 - C$ のとき，ETF に価格 $B_1 + B_2 - C$ の 1 株の買い注文を，$S_{ETF} > S_1 + S_2 + C$ のとき，ETF に価格 $S_1 + S_2 + C$ の 1 株の売り注文を出す．両方出す場合もあることに注意．その後，ETF に出していた買い注文が成立したら，直ちに株式 1 と 2 にそれぞれ価格 B_1, B_2 の売り注文を 1 株ずつ出す．これらの注文は対当する注文がすでに存在するのですぐに取引が成立し，裁定取引が完了する．ETF に出していた売り注文が成立したら，直ちに株式 1 と 2 にそれぞれ価格 S_1, S_2 の買い注文を 1 株ずつ出す．また，$B_1, B_2, B_{ETF}, S_1, S_2, S_{ETF}$ のいずれかが変更になれば，注文を入れなおす．

裁定エージェントのこれらの取引は必ず C だけの価格差を確保しており損失を出すことはないが，裁定取引の機会が全くない可能性はある．

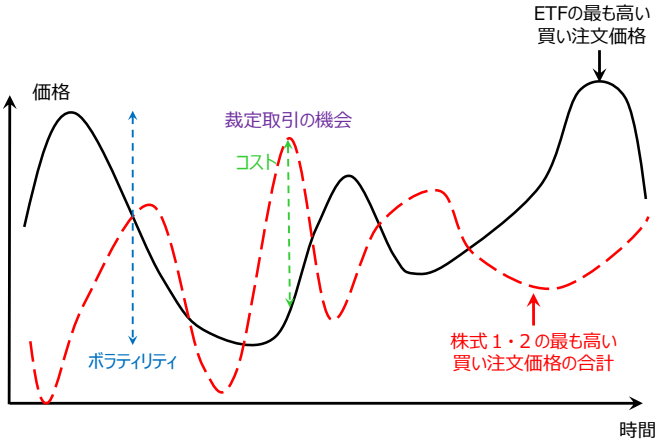


図 6: ボラティリティとコストの関係

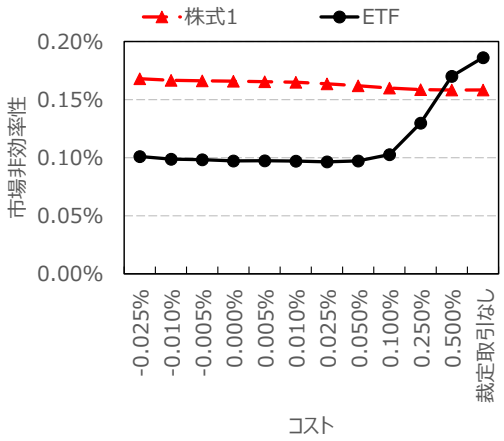


図 7: 市場非効率性 (Case A)

3. シミュレーション結果

本研究では [水田 13] と同じ以下のパラメータを用いる．具体的には，$n = 1000$, $w_{1,max} = 1$, $w_{2,max} = 10$, $w_{3,max} = 1$,$\tau_{max} = 10000$,$\sigma_\epsilon = 0.06$, $P_\sigma = 30$, $t_c = 20000$,$\delta P = 0.01, k = 0.1, P_{f0} = 10000$ とした．つまり各証券のファンダメンタル価格は，株式 1 ， 2 が $P_{f0} = 10000$, ETF が $2P_{f0} = 20000$ である．またシミュレーションは時刻 $t = t_e = 1000000$ まで行った [*8]． また，$c = -0.025\%, -0.01\%, -0.005\%, 0\%, 0.005\%, 0.01\%, 0.025\%, 0.05\%, 0.1\%, 0.5\%$, および裁定エージェントがいない場合に対して，その他の条件を乱数表も含め全く同じにして，各種統計値を算出した．これを，乱数表を変更して 100 回行い，その平均値を用いる．

Case A では，$k_0 = 5\%, k_1 = k_2 = ka = 50\%$ で固定する．Case B では表 1 で示されたようにコストごとに異なる，

$$k_1 = k_2 = k_{B,c} = k_A \times \frac{V_c}{V_0}, \tag{4}$$

とする．ここで，V_c は Case A でのコストごとの株式 1 の 20000 期あたり平均売買数量，V_0 は裁定取引エージェントが存在しないときの V_c である．表 1 に示されるように，コストが低いほど売買数量は多い．実際の市場では，売買の量が増えればより多くの量を注文する投資家がいることが知られている．この効果を取り入れるため，売買が増えた分，ノーマルエージェントは注文量を増やす．また，k_0 は Case A 同様，$k_0 = 5\%$ とする．

*8 これらのパラメータの妥当性検証については付録 “モデルの妥当性検証” 参照．また具体的なパラメータの検討は [水田 14] に書かれている．

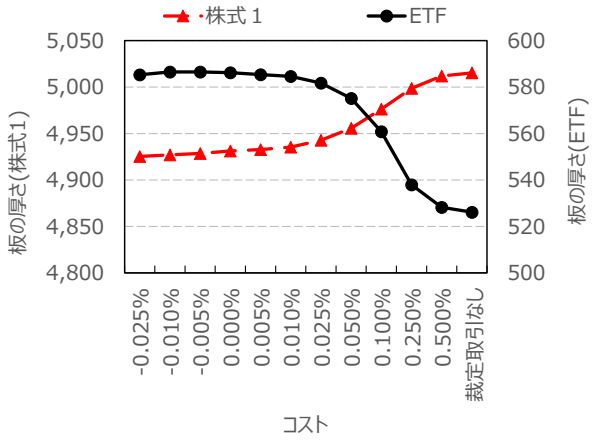


図 8: 板の厚さ (Case A)

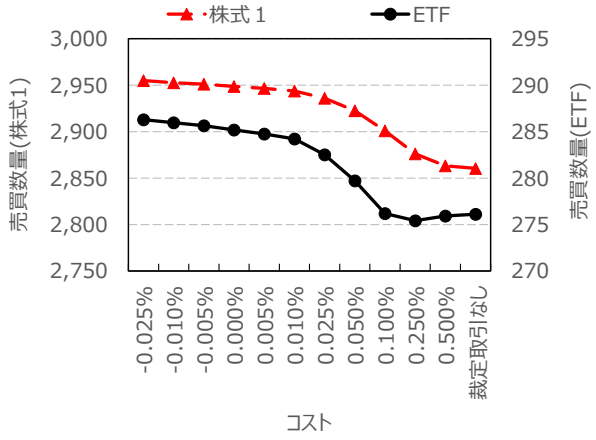


図 9: 売買数量 (Case A)

3.1 Case A： $k_1 = k_2 = 50\%$ で固定

図 5 はコスト c ごとの価格乖離率と裁定エージェントの売買株数である．価格乖離率は株式 1 と 2 の価格の合計と ETF の価格がどれだけ乖離していたかを示す指標であり，

$$M_d = \frac{1}{te} \sum_{t=1}^{te} \frac{|P_{ETF}^t - (P_1^t + P_2^t)|}{P_1^t + P_2^t} \tag{5}$$

と定義した．ここで，P_{ETF}^t, P_1^t, P_2^t はそれぞれ時刻 t における ETF，株式 1 ， 2 の価格，|| は絶対値である．

コストが減少すると裁定エージェントの売買が増え，価格の乖離が減っている．価格の乖離はコストが 0.1% 付近を閾値に急激に変化している．この値はちょうど，10 期ごとの株式のリターンの標準偏差（ボラティリティ）である 0.11% に近い値である．そのため，コストがボラティリティに比べ，高いか低いかが重要な境界であることが考えられる．

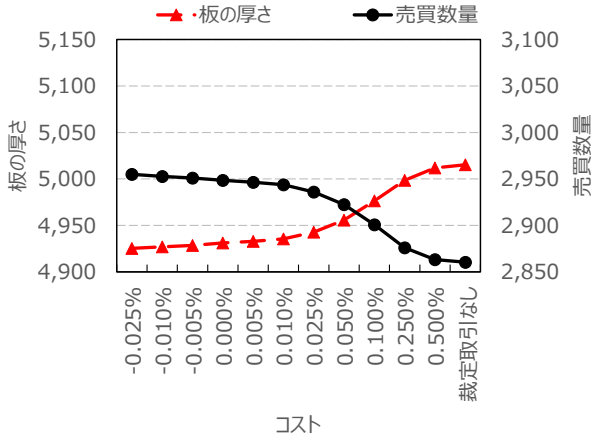


図 10: 株式 1 の板の厚さと売買数量 (Case A)

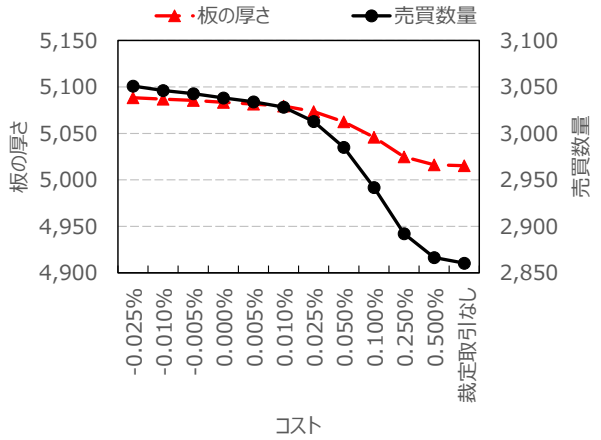


図 11: 株式 1 の板の厚さと売買数量 (Case B： 表 1 のように売買が増えた分 k_1, k_2 を増やす)

また，コストがマイナスになっても大きな変化が見られない．つまり，手数料をとられるかリベートを受け取れるかの違いは，コストがボラティリティに比べ高いか低いかに比べ，重要ではないことを示している．

図 6 は，ボラティリティとコストの関係を示した模式図である．赤い破線は株式 1， 2 の最も高い買い注文の価格の合計，黒い線は ETF の最も高い買い注文の価格である．そのため，赤い破線が黒い線よりコスト以上に上に来た場合のみ，裁定取引の機会がある．一方，各証券はボラティリティ程度の幅を持って価格が変動している．そのため，ボラティリティよりコストが小さければ裁定の機会が訪れやすく，裁定エージェントの売買が増え，ETF と株式の価格の乖離が小さくなる．

図 7 はコスト c ごとの ETF と株式 1 の市場非効率性である．市場の効率性を測定する指標として，市場非効率性 M_{ie}，

$$M_{ie} = \frac{1}{t_e} \sum_{t=1}^{t_e} \frac{|P^t - P_f|}{P_f} \tag{6}$$

を定義した [*9]．ここで $||$ は絶対値を示す．M_{ie} は 0 以上の値をとり，0 なら完全に効率的，大きくなればなるほど非効率であることを示す．

コストが減少すると ETF のみ市場が効率になっている．やはり，コストが 0.1% 付近を閾値に急激に変化している．株式 1 は非効率になったりはしていない．株式 1 の効率性を犠牲にして ETF を効率的にしているわけではないことが分かる．

図 8 は ETF と株式 1 の板の厚さ (待機している注文量) を示している．板の厚さは最も高い買い注文価格と最も安い売り注文価格の平均から ±0.1% の範囲にある注文株数の合計の全時間での平均である．ETF はコストが下がると板が厚くなっており，特にコスト 0.1% 付近で大きく変化している．株式 1 は逆の傾向となっている．図 9 はノーマルエージェントも含めた全売買株数である．ETF，株式 1 ともにコストが低くなると売買が増えている．コストの減少により株式 1 の板の厚さが減少し売買が増えていることは，株式 1 の待機注文が裁定取引と対当していると考えれば整合的である．

しかしながら，[Muranaga 99, Chung 05] などの多くの実証研究が売買量と板の厚さは正の相関があることを示しており，株式 1 の結果はこれと合わない． 図 10 は株式 1 の板の厚さと売買数量であるが，逆相関になっており実証研究の結果とあっていない．

3.2 Case B：表 1 で示された注文比率を k_1, k_2 に使用

実際の市場では，売買の量が増えればより多くの量を注文する投資家がいることが知られている．Case A ではこの効果を取り入れていなかった．この効果を取り入れるため，Case B では売買が増えた分，ノーマルエージェントは注文量を増やす．

図 11 は Case B での，株式 1 の板の厚さと売買数量を示した．コストが減少するにつれて，板の厚さも売買数量も増加していて，実証研究とあっている．

このことから，売買の量が増えればより多くの量を注文する投資家の存在が，流動性の統計的性質を決める重要なメカニズムである可能性を示している．

4. まとめと今後の課題

本研究では，[水田 19a] の人工市場モデルに，売買数量を調整できる仕組みを実装して実証分析と整合的な結果を出力するモデルを構築し，[水田 19a] では調べられていなかったリベートも含め，裁定取引にかかるコストが流動性に与える影響を分析した．

その結果，ボラティリティよりコストが小さければ裁定の機会が訪れやすく，裁定エージェントの売買が増え，ETF と株式の価格の乖離が小さくなることが分かった．また，コストがマイナスになっても大きな変化が見られなかった．このことは，手数料をとられるかリベートを受け取れるかの違いは，コストがボラティリティに比べ高いか低いかに比べ，重要ではないことを示している．

さらに，売買数量にかかわらず注文数量を一定にした場合は，株式の売買量と板の厚さは負の相関となり実証研究とあわなかった．一方で，売買数量の増加に応じて注文数量を増やした場合は，株式の売買量と板の厚さは正の相関となり実証研究と整合的な結果が得られた．このことは，売買の量が増えればより多くの量を注文する投資家の存在が，流動性の統計的性質を決める重要なメカニズムである可能性を示している．

*9 市場の効率性を示す指標は多く提案されている [伊藤 07, Verheyden 13] が，本研究で用いる市場非効率性は，通常は観測できないファンダメンタル価格 P_f を直接使用しており，人工市場シミュレーションでしか用いることができない定義である．人工市場シミュレーションでは P_f が明確であるため，推定ではない理想的な市場の効率性を測定できる市場非効率性の使用が可能である．

本研究では，裁定取引へのリベートしか分析しなかった．リベートは高頻度取引を増やして流動性を向上させると言われているが，高頻度取引には裁定取引以外にも，マーケットメイク戦略などさまざま存在する．これらの戦略へのリベートの効果は今後の課題である．

また，リベートが高くなると，テイク側の手数料も高くならざるを得ないが，テイク側にとって流動向上のメリットと手数料が高くなるデメリットのどちらが大きいのか検討する必要がある．このような観点で分析していないため，今後の課題である．

付録

モデル構築の基本理念

人工市場シミュレーションを用いれば，まだ導入したことがない手数料体系を調べたり，その変更の純粋な効果やメカニズムを議論したりできる．これが人工市場シミュレーション研究の強みである． そして，多くの人工市場シミュレーション研究がこれまでにない環境が与える影響や，規制・制度の変更を分析してきた [水田 19b, Mizuta 19c].

ただその効果は確実な予想ではない．さまざまなケースでのシミュレーションを行い，これまで予想されていなかった，“あり得る” メカニズムでの現象を見つけておくことが，人工市場シミュレーションの大きな役割となる．金融市場でこれから実際におこる現象を定量的にも忠実に再現することが目的ではなく，環境の変化が，どのようなメカニズムで価格形成に影響を与え，どのようなことが起こり得るのかという知識獲得が目的である．これは例えば実証分析など他の手法ではできないことである．

人工市場モデルは普遍的に存在するマクロ現象を再現すべきであると考えられる．人工市場シミュレーションでは，マクロ現象である市場価格のリターンや売買数量をモデル化しない．あくまで，投資家を模した “エージェント” と取引所を模した “価格決定メカニズム” といったミクロメカニズムをモデル化し，そのミクロメカニズムの相互作用の積み上げとしてマクロ現象が出力される．そのため，ミクロメカニズムのモデル化は現実の市場に即したものとし，結果として出力されるマクロ現象は，現実の市場で普遍的に存在するマクロ的性質を再現されるように作る必要がある．

しかし，普遍的ではなく特定の時期や資産，地域で出現するマクロ的性質すべてを再現することは本研究の目的ではない．必要以上に多くのマクロ的性質を一つのモデルで再現しようとすると，過剰に複雑なモデルをもたらし，関連する要素が多くなりすぎて，発生メカニズムの理解を妨げてしまう．

実際，複雑な人工市場モデルに対して，モデルが複雑になるとパラメータが増えモデルの評価が困難になるという批判がある [Chen 12]．モデルが複雑すぎると関連する要素が多くなりすぎて，発生メカニズムの理解を妨げてしまう．また, パラメータが増えるほどさまざまな出力がだせるようになり，モデルを作った人が導きたい結果へ恣意的に導くためのパラメータ設定が行われる恐れがある．シンプルでパラメータが少ないモデルほど，パラメータ調整によって特定の結果に導くことが困難であるため評価が容易となる．

以上により，本研究では，分析目的を果たせる範囲内でなるべくシンプルなモデルの構築を行っている．実際の市場を完全に再現することを目的としておらず，普遍的ではなく特定の時期や資産，地域で出現するマクロ的性質すべてを再現することや, 実際には存在するであろう投資家をすべて網羅することはあえて行っていない．

表 2: 裁定エージェントがいない場合の株式 1 の各種統計量

取引	約定率	32.3%
	キャンセル率	26.1%
	注文件数 / 20000 期間	6467
標準偏差	1 期間	0.0512%
	20000 期間	0.562%
	尖度	1.42
二乗リターンの自己相関係数	ラグ	
	1	0.225
	2	0.138
	3	0.106
	4	0.087
	5	0.075

[Weisberg 12] が述べているように，よいシミュレーションモデルとはその研究目的によって異なる．そのため，本研究のモデルは本研究の目的にのみおいてよいモデルであり，他の研究目的においてはよいモデルではない．また [Weisberg 12] が述べているように，数理モデルと異なり，シミュレーションモデルは投資家などのミクロの行動やその行動の理由と，市場価格などのマクロ現象との相互作用のメカニズムの解明ができることが強みである．例えば，ノーベル経済学賞を受賞したシェリングはシミュレーションモデルを用いて人種隔離のメカニズムを議論したが，そのモデルではメカニズムの解明に焦点を当てるためどこにも存在しないシンプルな町を簡潔にモデル化していて，実際に存在する町の隔離状況の説明や予測はできないモデルであった [Schelling 06]．本研究でも，数理モデルが強みとするマクロ現象の特徴分析や予測といったことは目的とせず，メカニズムの解明に焦点をあてている．

モデルの妥当性

人工市場モデルの妥当性は実証分析で得られている fat-tail や volatility-clustering といった代表的な stylized fact が再現できるかどうかで評価される [LeBaron 06, Chen 12, 水田 14]. ファット・テールは，市場価格のリターンの分布が正規分布ではなく裾が厚い，すなわち，尖度が正であることである．ボラティリティ・クラスタリングは市場価格のリターンの 2 乗が，大きなラグでも自己相関係数が有意に正であることである．

[Sewell 11] など多くの研究で述べられているように，金融市場は不安定であり，安定的に，どのような時期にも有意に観測されるスタイライズド・ファクトはファット・テールとボラティリティ・クラスタリングの 2 つしかない．

しかも，これらは統計量の有意に正であることだけが安定して観測され，値そのものは，時期によって異なる．ファット・テールについては，実証分析でよく観測されるリターンの分布の尖度は 1 ～ 100 程度であり，ボラティリティ・クラスタリングについては，実証分析でよく観測されるリターンの自己相関は 0 ～ 0.2 程度と，かなりばらつきがある [Sewell 11].

本研究のように，金融市場に共通する性質を分析対象とする人工市場が再現すべきは，これらの統計量が有意に正であり，問題ない範囲に値が収まっていることであって，特定の値に近づけることは本質的ではない．

表 2 は，裁定エージェントがいない場合の株式 1 の毎期のリターン $\log(P^t/P^{t-1})$ の標準偏差と尖度，リターンの 2 乗の自己相関である．リターンの尖度がプラスで，ファット・テー

ルが再現されている．また，リターンの 2 乗の自己相関もプラスで，ボラティリティ・クラスタリングが再現されていると考えられる．

留意事項

本論文はスパークス・アセット・マネジメント株式会社の公式見解を表すものではありません．すべては個人的見解であります．

参考文献

[Battiston 16] Battiston, S., Farmer, J. D., Flache, A., Garlaschelli, D., Haldane, A. G., Heesterbeek, H., Hommes, C., Jaeger, C., May, R., and Scheffer, M.: Complexity theory and financial regulation, *Science*, Vol. 351, No. 6275, pp. 818–819 (2016),
http://science.sciencemag.org/content/351/6275/818

[Bookstaber 17] Bookstaber, R.: *The End of Theory: Financial Crises, the Failure of Economics, and the Sweep of Human Interaction*, Princeton University Press (2017), (監修： 長尾慎太郎，邦訳: 井田京子: 経済理論の終焉金融危機はこうして起こる，Pan Rolling(2019))
http://www.panrolling.com/books/wb/wb273.html

[Chen 12] Chen, S.-H., Chang, C.-L., and Du, Y.-R.: Agent-based economic models and econometrics, *Knowledge Engineering Review*, Vol. 27, No. 2, pp. 187–219 (2012),
http://dx.doi.org/10.1017/S0269888912000136

[Chiarella 02] Chiarella, C. and Iori, G.: A simulation analysis of the microstructure of double auction markets, *Quantitative Finance*, Vol. 2, No. 5, pp. 346–353 (2002),
https://doi.org/10.1088/1469-7688/2/5/303

[Chung 05] Chung, K. H., Kim, K. A., and Kitsabunnarat, P.: Liquidity and Quote Clustering in a Market with Multiple Tick Sizes, *Journal of Financial Research*, Vol. 28, No. 2, pp. 177–195 (2005),
https://doi.org/10.1111/j.1475-6803.2005.00120.x

[Farmer 09] Farmer, J. D. and Foley, D.: The economy needs agent-based modelling, *Nature*, Vol. 460, No. 7256, pp. 685–686 (2009),
https://www.nature.com/articles/460685a

[伊藤 07] 伊藤 幹夫：効率的市場仮説をめぐる論争はなぜ決着しないのか, 三田学会雑誌 (慶應義塾大学), Vol. 100, No. 3, p. 229 (2007),
http://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-20071001-0211

[和泉 17a] 和泉 潔, 川久保 佐記, 米納 弘渡：第 5 章 強靭な金融システム, 古田 一雄（編）, レジリエンス工学入門, 日科技連出版社 (2017),
http://www.juse-p.co.jp/cgi-bin/html.pl5?i=ISBN978-4-8171-9624-8

[和泉 17b] 和泉 潔：第 6 章 可能世界ブラウザとしてのエージェントシミュレーション，マルチエージェントのためのデータ解析 (マルチエージェントシリーズ)，コロナ社 (2017),
http://www.coronasha.co.jp/np/isbn/9784339028126/

[川本 15] 川本 隆雄：米国における株式市場構造改革議論とその行方, 月刊資本市場, Vol. 11, No. 363 (2015),
http://www.camri.or.jp/files/libs/335/201703251704004867.pdf

[LeBaron 06] LeBaron, B.: Agent-based computational finance, *Handbook of computational economics*, Vol. 2, pp. 1187–1233 (2006),
http://dx.doi.org/10.1016/S1574-0021(05)02024-1

[水田 13] 水田 孝信, 早川 聡, 和泉 潔, 吉村 忍：人工市場シミュレーションを用いた取引市場間におけるティックサイズと取引量の関係性分析, JPX ワーキング・ペーパー, No. 2, 日本取引所グループ (2013),
https://www.jpx.co.jp/corporate/research-study/working-paper/index.html

[水田 14] 水田 孝信：人工市場シミュレーションを用いた金融市場の規制・制度の分析, 博士論文，東京大大学大学院工学系研究科 (2014),
http://hdl.handle.net/2261/59875

[Mizuta 16] Mizuta, T.: A Brief Review of Recent Artificial Market Simulation (Agent-Based Model) Studies for Financial Market Regulations and/or Rules, *SSRN Working Paper Series* (2016),
http://ssrn.com/abstract=2710495

[水田 18] 水田 孝信：なぜ株式市場は存在するのか？, スパークス・アセット・マネジメント (2018),
https://www.sparx.co.jp/report/special/2174.html

[水田 19a] 水田 孝信：株式と ETF の裁定取引にかかるコストと流動性の関係－人工市場によるシミュレーション分析－, JPX ワーキング・ペーパー, No. 27, 日本取引所グループ (2019),
https://www.jpx.co.jp/corporate/research-study/working-paper/index.html

[水田 19b] 水田 孝信：人工市場シミュレーションを用いた金融市場の規制やルールの議論, 証券アナリストジャーナル, Vol. 57, No. 5 (2019),
https://www.saa.or.jp/learning/journal/each_title/2019/05.html

[Mizuta 19c] Mizuta, T.: An agent-based model for designing a financial market that works well, *arXiv* (2019),
https://arxiv.org/abs/1906.06000

[Muranaga 99] Muranaga, J.: Dynamics of Market Liquidity of Japanese Stocks: An Analysis of Tick-by-Tick Data of the Tokyo Stock Exchange, in International Settlements, for B. ed., *Market Liquidity: Research Findings and Selected Policy Implications*, Vol. 11 of *CGFS Papers chapters*, pp. 1–25, Bank for International Settlements

(1999),
https://www.bis.org/publ/cgfs11.htm

[岡田 17] 岡田 功太, 齋藤 芳充 : 米国株式市場のメイカー・テイカー・モデルを巡る議論-流動性向上策としてのリベートの功罪-, 野村資本市場クォータリー, Autumn (2017),
http://www.nicmr.com/nicmr/report/repo/2017/2017aut04.pdf

[Schelling 06] Schelling, T. C.: *Micromotives and Macrobehavior*, W. W. Norton & Company (2006), (邦訳: 村井章子 : ミクロ動機とマクロ行動，勁草書房 (2016))
http://www.keisoshobo.co.jp/book/b251669.html

[Sewell 11] Sewell, M.: Characterization of financial time series, *Research Note, University College London, Department of Computer Science*, No. RN/11/01 (2011),
http://finance.martinsewell.com/stylized-facts/

[Todd 16] Todd, A., Beling, P., Scherer, W., and Yang, S. Y.: Agent-based financial markets: A review of the methodology and domain, in *Computational Intelligence for Financial Engineering Economics (CIFEr), 2016 IEEE Symposium Series on Computational Intelligence on* (2016),
https://doi.org/10.1109/SSCI.2016.7850016

[東証 15] 東証 : 東証公式 株式サポーター 株式取引編, 東京証券取引所 (2015),
https://www.jpx.co.jp/learning/tour/books-brochures/detail/08.html

[Verheyden 13] Verheyden, T., De Moor, L., and Bossche, Van den F.: A Tale of Market Efficiency, *Review of Business and Economic Literature*, Vol. 58, No. 2, pp. 140–158 (2013),
https://ssrn.com/abstract=2382793

[Weisberg 12] Weisberg, M.: *Simulation and Similarity: Using Models to Understand the World*, Oxford Studies in the Philosophy of Science (2012), (邦訳: 松王政浩: 科学とモデル-シミュレーションの哲学 入門-，名古屋大学出版会 (2017))
http://www.unp.or.jp/ISBN/ISBN978-4-8158-0872-3.html

人工市場を用いた株式貸借市場の流動性の変化が売買市場に与える影響の分析

Analysis of the impact of changes in the stock lending market liquidity on the stock market using an artificial market

宮崎 文吾[1*]　中川 慧[1]
Bungo Miyazaki[1] Kei Nakagawa[1]

[1] 野村アセットマネジメント株式会社
[1] Nomura Asset Management Co.,Ltd.

Abstract: 株式貸借市場は，空売りを通じて株式市場（売買市場）に流動性を提供するとともに，資金調達機会や収益機会を提供するという重要な役割を担っている．空売りやその規制が株式市場にどのような影響を与えるかについて多くの研究が存在しており，人工市場はその分析手法の一つであるが，既存の人工市場を用いた研究は現実の空売りのメカニズムを適切に反映できていない．つまり，貸借市場における供給制約や需給に基づく貸株料も空売りへの制約となり，効率的な株式市場の形成のためには流動性の高い貸借市場の存在が条件となると考えられる．そこで本研究では株式市場に加え，貸借市場も考慮した連成型人工市場を構築し，貸借市場の流動性の変化が株式市場の効率性に与える影響を調べる．分析の結果，貸借市場の流動性が高まると平常時の売買市場の効率性は高まるものの，ファンダメンタル価格が急落した際の効率性は損なわれる可能性があることが分かり，空売り規制の設計の際などに貸借市場の流動性も考慮する必要があることを示した．

1　はじめに

株式貸借市場は，空売りを通じて株式市場（取引所売買市場，以下「売買市場」）に流動性を提供するとともに，資金調達機会や収益機会を提供するという重要な役割を担っている．すなわち，空売りを行う投資家にとっては，ネイキッド・ショート・セリングが原則的に禁止されている本邦において貸借市場で株式を借り入れることで決済措置をとることができ，一方で貸株を行う投資家にとっては，短期間で売却する予定のない株式を担保とした資金調達を行ったり，貸株料によって運用効率の向上を図れる（図 1）．
空売りやその規制が売買市場にどのような影響を与えるかについて多くの研究が存在しており，[1, 2] などをはじめとして人工市場はその分析手法の一つである．

ところが，既存の人工市場を用いた研究のほとんどは貸借市場での「流動性」の存在を考慮していない．本研究では貸借市場における流動性を表す尺度として，(1) 貸株供給量と (2) 貸株料の 2 つを対象とする．これらは現実の空売りのメカニズムで考慮が必要な要素であり，空売りに対する自然な制約として存在している．具体的には，貸借市場での貸株供給量が少なく，空売りするための株式の調達が困難な場合や，貸株料が空売りを行うことでの期待リターンに見合うほど安くない場合に，空売りの制約となる．これらは貸借市場での需給に基づき発生する制約であり，裁定取引に対する制約条件 (裁定の限界) として存在する [3]．したがって，市場分析の際には現実の空売りのメカニズムを適切に反映するように明示的にモデル化する必要があり，このような需給に基づく制約を考慮せず空売り規制などの制度設計をしてしまうと，取引参加者に対して過剰な制約を課し市場の効率性を損う恐れがある．
そこで本研究では株式の売買市場に加え，上記の流動性の制約を持った貸借市場を考慮した連成型人工市場を構築する．そして，貸借市場の流動性の変化，特に貸株料の変化が売買市場の効率性に与える影響を調べる．

*連絡先：野村アセットマネジメント株式会社
〒 103-0027 東京都中央区日本橋 1 丁目 12-1
E-mail: b-miyazaki@nomura-am.co.jp

2　先行研究

貸借市場と売買市場の関係を調査した実証研究としては，[4, 5] などが存在する．これらの研究ではいずれも，貸借市場の流動性（前者では貸株在庫残高に対する貸株残高や貸株料，後者では時価総額に対する貸株在庫残高や貸株料などを指標としている）が売買市場の流動性や効率性（前者では取引件数やビッドアスク

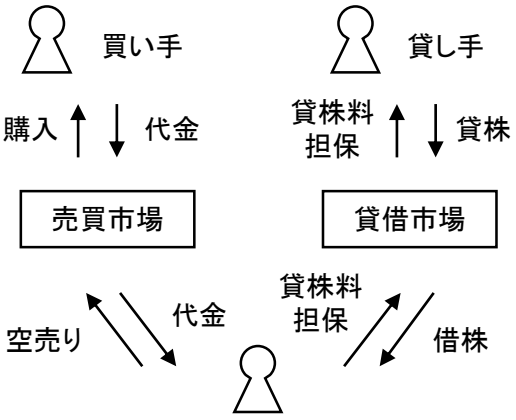


図 1: 売買市場と貸借市場の関係

スプレッド，後者では市場リターンが株式リターンへ織り込まれる速度などを指標としている）に影響を与えることを示している．
本研究において売買市場と貸借市場を考慮するように，複数の関連しあった市場を考慮した人工市場を連成型人工市場といい，例えば原資産市場とオプション市場を調べた [6] や現物と ETF との裁定取引を調べた [7] などが存在する．また，空売り規制の影響を調査するためのエージェントシミュレーションとして，[8] は貸借市場での供給制約や貸株料を考慮した戦略を持ったエージェントを再現した．しかし，この研究での貸株料は，ある値を基準として，過去の市場価格が上昇している場合は一様乱数から定まる値だけ下落，市場価格下落している場合は上昇というように定められ，現実のような貸借市場での需給を反映したモデルとなっておらず，貸借市場の流動性の変化が売買市場に与える影響の調査という目的では不十分である．

3 売買市場と貸借市場の連成型人工市場モデルの構築

本研究では [9] をベースとした [2] の人工市場モデルをベースに，貸借市場を導入したモデルを構築する．本章ではまず，[2] でのモデルについて述べ，そのモデルへの貸借市場の導入を行う．

3.1 売買市場のモデル

はじめに売買市場を，[2] と同様に構築する（以下「既存モデル」と呼ぶ）．既存モデルでは，一つの資産を対象とした複数のエージェントによる連続ダブルオークションをモデル化しており，n 体のエージェントが以下のプロセスに基づき順番に注文を行うことを繰り返す．

3.1.1 注文の決定

ある時刻 t でエージェント j は次のステップに基づき注文を行う．なお，既存モデルではエージェントが 1 体注文を出すごとに時刻 t は 1 増えるとしたが，本研究では全エージェントが注文を出した後に 1 増えるものとした．

Step 1. 期待リターンの計算

あるエージェント j が時刻 t において注文を行おうとしている時点を t' で表す．期待リターン $r^{t'}_{e,j}$ を次式のように，ファンダメンタル価格と直近約定価格の乖離 $\log(P^t_f/P^{t'})$(ファンダメンタル成分)，ヒストリカルリターン $r^{t'}_{h,j}$(テクニカル成分)，ノイズ $\varepsilon^{t'}_j$ の加重平均で計算する．

$$r^t_{e,j} = \frac{\left(w_{1,j}\log\frac{P^{t'}_f}{P^{t'}} + w_{2,j}r^{t'}_{h,j} + w_{3,j}\varepsilon^t_j\right)}{\sum_{i=1}^{3} w_{i,j}} \quad (1)$$

$$w_{i,j} \in [0, w_{i,max}] \quad (2)$$

$$\varepsilon^{t'}_j \sim N(0, \sigma_\varepsilon) \quad (3)$$

$$P^t_f : \text{ファンダメンタル価格} \quad (4)$$

$$P^{t'} : \text{直近約定価格} \quad (5)$$

$$r^{t'}_{h,j} = \log(P^{t'}/P^{t-\tau_j}) \quad (6)$$

$$\tau_j \in [1, \tau_{max}] \quad (7)$$

なお，各要素の重み $w_{i,j}$ および過去リターンの参照期間 τ_j は，エージェントごとにシミュレーション開始時に一様乱数を用いて生成する．また，空売りを意図的に発生させるために，本研究では各エージェントはエージェントに固有のバイアス b_j をもってファンダメンタル価格を認識するものとし，式 (1) におけるファンダメンタル価格 P^t_f の代わりにエージェント j が認識するバイアス付きのファンダメンタル価格 $P^t_{f,j} = P^t_f \exp b_j$ を用いた．$b_j < 0$ となるエージェントは実際のファンダメンタル価格より悲観的にファンダメンタル価格を認識するため，低い価格での売り注文が増え，空売りを行いやすくなる．b_j はシミュレーション開始時に $b_j \sim U[-b_{max}, b_{max}]$ で初期化した．

Step 2. 注文価格の決定

期待リターン $r^{t'}_{e,j}$ をもとに，期待価格 $P^{t'}_{e,j}$ の計算を行う．

$$P^{t'}_{e,j} = P^{t'} \exp(r^{t'}_{e,j}) \quad (8)$$

次に，注文価格 $P_{o,j}^{t'}$ を以下の通り決定する．

$$P_{o,j}^{t'} = P_{e,j}^{t'} \exp(r_d^{t'}) \tag{9}$$

$$r_d^{t'} \sim U[-r_{max}, r_{max}] \tag{10}$$

ここで，r_{max} は期待価格と注文価格の差の最大幅を表す定数である．最後に，$P_{o,j}^{t'}$ を価格の変化幅の最小単位 δP に収まるように丸める．なお，既存モデルでは式 (9) の代わりに $P_{o,j}^{t'}$ は $P_{e,j}^{t'}$ から一定幅内の一様乱数で選択されるとしたが，本研究では期待価格の絶対的な値によらずに注文価格を決められるように，このように選択するものとした．また既存モデルでは最小単位 δP への丸めの際に切り捨てとしたが，本研究では δP の整数倍のうち $P_{o,j}^{t'}$ からより近い方に丸めた．

Step 3. 注文

注文価格 $P_{o,j}^{t'}$ と期待価格 $P_{e,j}^{t'}$ に基づき，次のように注文を行う．

$$1 \text{ 単位の買い注文 } (P_{e,j}^{t'} > P_{o,j}^{t'}) \tag{11}$$

$$1 \text{ 単位の売り注文 } (P_{e,j}^{t'} < P_{o,j}^{t'}) \tag{12}$$

マッチングする注文が板に存在する場合は即座に相手方の価格で約定し直近価格を更新する．そのような注文がなければ板に注文を残す．板に残された注文が t_c 経過しても約定されない場合，その注文はキャンセルされる．

3.1.2 エージェントの学習

各エージェントは自身の注文を出す前に以下の手順で戦略の更新を行う．式 (1) におけるファンダメンタルおよびテクニカルに対応する期待リターンの成分をそれぞれ，

$$r_{e,i,j}^{t'} = \begin{cases} \log(P_f^{t'}/P^{t'}), & i = 1 \\ r_{h,j}^{t'}, & i = 2 \end{cases} \tag{13}$$

とする．価格変化の方向を当てている戦略のウエイトを引き上げ，外れている戦略のウエイトを引き下げる学習を行う． すなわち，ファンダメンタル，テクニカルに対応するウェイト $w_{i,j}$ を $r_{e,i,j}^{t'}, i = 1, 2$ と学習期間のリターン $r_l^{t'} = \log(P^{t'}/P^{t-t_l})$ のそれぞれの符号を比べて次式のように調整する．

$$w_{i,j}^{new} = \begin{cases} w_{i,j}^{old} + k_l r_l^{t'} \rho_j^t (w_{i,max} - w_{i,j}^{old}) & \text{(同符号の時)} \\ w_{i,j}^{old} - k_l r_l^{t'} \rho_j^{t'} w_{i,j}^{old} & \text{(異符号の時)} \end{cases} \tag{14}$$

$$k_l : \text{定数} \tag{15}$$

$$\rho_j^{t'} \sim U[0, 1] \tag{16}$$

ここで，k_l は定数，$\rho_j^{t'}$ は時刻ごと，エージェントごとに与えられる 0 から 1 までの一様乱数である．
また，これとは別に各エージェントは小さい確率 m で $w_{i,j}, b_j$ をシミュレーション開始時と同じ一様乱数に従い再設定する．

3.2 貸借市場の導入

つぎに，本研究で導入する貸借市場とそれに対する各エージェントの振る舞いについて説明する．あるエージェントの売り注文が約定する際，それが空売りである場合（自己の保有する株式が 0 以下の場合），エージェントは貸借市場に対して貸株供給量と貸株料を確認し，条件を満たす場合のみ株式を借り入れて空売りを行い，満たさない場合はその注文はキャンセルされる．

したがって，本モデルでのエージェントは貸借市場での流動性の尺度である (1) 貸株供給量や (2) 貸株料といった制約の存在を考慮しているため，現実の投資家の意思決定をより適切に反映していると考えられる．

具体的には，以下のステップに基づき空売りの可否を判断する．

Step 1. 総需要量の計算

貸借市場で全エージェントが借り入れている株式の合計である需要量 $D^{t'}$ を計算する．すなわち，

$$D^{t'} = \sum_{j=1}^{n} si_j^{t'} \tag{17}$$

とする．ここで，$si_j^{t'}$ は t' におけるエージェント j の空売り中の株式であり，$q_j^{t'}$ を各エージェントの各時点での保有株数（マイナスを許す）として，

$$si_j^{t'} = \max(0, -q_j^{t'}) \tag{18}$$

で表される．

Step 2. 総供給量の計算

次に，貸借市場にて各エージェントが貸し出すことのできる株式数の合計である供給量 $S^{t'}$ を計算する．すなわち，

$$S^{t'} = \sum_{j=1}^{n} li_j^{t'} \tag{19}$$

とする．ここで $li_j^{t'}$ は t' におけるエージェント j の貸し出し可能株数であり，

$$li_j^{t'} = \max(0, \lfloor \max(0, q_j^{t'}) * lr_j \rfloor - us_j^{t'}) \tag{20}$$

とする．ここで，$\lfloor x \rfloor$ は x 以下の最大の整数を表す記号，lr_j はエージェント j の保有株数に対し

て貸し出し可能な株数の比率を表す $[0,1]$ 内の実数，$us_j^{t'}$ は時刻 t' におけるエージェント j の未約定の売り注文の合計である．

Step 3. 貸株料の決定
上述の $D^{t'}, S^{t'}$ を基に単位時間当たりの貸株料 $f^{t'}$ が次式で決定される．

$$f^{t'} = \alpha \exp(\beta \frac{D^{t'}+1}{S^{t'}}) \tag{21}$$

ここで，α, β はそれぞれ貸借市場の流動性を表すパラメータであり，α は全体的な貸株料の高低を表すパラメータ，β は貸株需給の変化に対する貸株料の感応度を表すパラメータである．それぞれ小さいほうが貸株料が低くなり，貸借市場の流動性が高いと言える．

Step 4. 空売りの判断
空売りを行おうとした際に，各エージェントは空売り注文が約定した際の期待リターンと貸株料を比較し，空売りを実行するかの判断を行う．すなわち，

$$P_{o,j}^{t'}/P_{e,j}^{t'} - 1 > f^{t'}\tau_j \tag{22}$$

を満たす際に空売りは実行され，満たさない場合はその注文はキャンセルされる．なお，キャンセルされる注文が新規に発注された売り注文であれば，ステップは次のエージェントの注文に移行し，新規の注文が買い注文で板に残っていた注文が空売りとなってキャンセルされる場合には，その売り注文をキャンセルし，次に約定の優先順位が高い売り注文と約定可能であるか判定する．

4 シミュレーション結果

本研究では，売買市場の各パラメータを以下の通り設定した．

エージェント数: $n = 1000$

ウェイト: $w_{1,max} = 1, w_{2,max} = 10, w_{3,max} = 1$

ノイズの分散: $\sigma_\varepsilon = 0.03$

過去リターンの参照期間: $\tau_{max} = 10$

バイアス: $b_{max} = 0.05$

機体価格と注文価格の差: $r_{max} = 0.1$

売買価格単位: $\delta P = 1$

注文キャンセル期間: $t_c = 10$

学習パラメータ: $t_l = 10, k_l = 4, m = 0.01$

貸借市場に関するパラメータについては，各エージェントの貸株比率を表す lr_j は全エージェントで共通とし，$lr_j = 0.8$ とした．

以上の設定のもとで流動性を表すパラメータ α, β を変化させ，貸借市場の流動性の変化が売買市場に与える影響を調べる．
各エージェントには，初期化の際に初期保有株数 q_j^0 を $U[0,20]$ で割り当てる．シミュレーションは $t = 6,000$ まで行い，状態が安定する $t = 1,000$ 以降の価格を調べた．

4.1 ファンダメンタル価格一定の場合

まず，平常時としてファンダメンタル価格が一定の $P_f^t = 10,000$ としてシミュレーションを行い，モデルの妥当性と貸借市場の流動性と売買市場の効率性の関係を調べた．モデルの妥当性の検証には，[2] と同様に現実の市場で見られる fat-tail や volatility clustering といった代表的な stylized fact を再現しているかで評価する．表 1 に，リターンの尖度，2 乗の自己相関関数について，それぞれのパラメータで 50 回シミュレーションを行った平均を示す．いずれのパラメータでも，尖度，自己相関関数共に有意にプラスとなっており，既存モデルに貸借市場を導入してもこれらの性質が保たれることが分かった．

表 1: リターンの尖度とリターンの 2 乗の自己相関関数

	α	0.00001					
	β	10	20	30	40	50	100
尖度		0.180**	0.190**	0.182**	0.187**	0.191**	0.202**
ACF	1	0.106**	0.105**	0.102**	0.109**	0.102**	0.104**
	2	0.036**	0.032**	0.034**	0.034**	0.029**	0.032**
	3	0.015**	0.015**	0.017**	0.017**	0.015**	0.019**
	4	0.009**	0.006**	0.002	0.006**	0.008**	0.006**
	5	0.000	-0.003	0.001	-0.001	-0.001	-0.001
	α	0.0001					
	β	10	20	30	40	50	100
尖度		0.178**	0.192**	0.190**	0.196**	0.196**	0.192**
ACF	1	0.102**	0.103**	0.107**	0.100**	0.109**	0.104**
	2	0.035**	0.035**	0.033**	0.032**	0.035**	0.033**
	3	0.012**	0.019**	0.015**	0.016**	0.018**	0.016**
	4	0.004*	0.007**	0.005**	0.005*	0.008**	0.006
	5	-0.001	-0.000	-0.004*	-0.002	0.001	0.000
	α	0.001					
	β	10	20	30	40	50	100
尖度		0.192**	0.182**	0.205**	0.208**	0.209**	0.216**
ACF	1	0.107**	0.101**	0.104**	0.102**	0.100**	0.103**
	2	0.032**	0.030**	0.033**	0.028**	0.033**	0.035**
	3	0.017**	0.017**	0.015**	0.015**	0.017**	0.019**
	4	0.005*	0.005*	0.007**	0.007**	0.004	0.006**
	5	0.002	-0.001	-0.000	0.001	-0.001	0.003

$*p < .05$, $**p < .01$

次に，貸借市場の流動性と売買市場の効率性の関係を検証する．売買市場の効率性の指標としては，取引価格とファンダメンタル価格の乖離の平均と乖離の絶対値の平均を用いた．表 2 にその結果をまとめる．いずれの場合でも，空売りの制約があることにより取引価格は平均的にファンダメンタル価格から上方に乖離し，貸借市場の流動性を表すパラメータ α, β が大きくなるほど（流動性が低くなるほど）その乖離が大きくなることが確認できた．

表 2: 取引価格とファンダメンタル価格の乖離の平均（カッコ内は乖離の絶対値の平均）

		β					
		10	20	30	40	50	100
α	1e-05	40.43 (48.71)	51.28 (56.20)	55.42 (59.59)	57.81 (61.61)	59.56 (63.03)	62.80 (65.93)
	0.0001	50.18 (55.22)	56.63 (60.57)	59.78 (63.15)	60.95 (64.16)	61.94 (64.98)	64.18 (66.91)
	0.001	59.51 (62.69)	62.41 (65.35)	63.65 (66.43)	64.16 (66.96)	64.60 (67.33)	65.52 (68.12)

4.2 ファンダメンタル価格が急落する場合

次に，[2] と同様に $t = 3,000$ でファンダメンタル価格を 10,000 から 7,000 に急落させた場合について調べた．図 2 にこの時の価格推移の例を示す．α, β が小さく貸借市場の流動性が高い状態では，空売り制約がほとんどない場合に相当する．この場合には，取引価格はファンダメンタル価格の急落後にオーバーシュート（ファンダメンタル価格を超えて下落）して，振動をしながらファンダメンタル価格に収束している．これは [2] で示されたような規制なしの売買市場と同様の特徴である．一方，α, β が大きく貸借市場の流動性が低い状態では，空売り制約が強い状態に相当する．この場合の取引価格推移はオーバーシュートすることなくファンダメンタル価格に近づいていき，[2] で示されたような完全空売り規制やアップティックルールの存在する売買市場と同様の特徴である．図 3, 4 を見ると，ファンダメンタル価格急落時に貸株残高が増えることで貸株料が上昇し，それに伴い $t = 3,050$ 前後で空売りの約定が成立していないことが分かる．このように，貸借市場での需給変動から生じる貸株料の上昇が，急落時に空売りへの制約を自然に強めていることがシミュレーションからも示せた．

貸借市場の流動性とオーバーシュートの大きさの関係を調べるため，表 3 にファンダメンタル急落以後の取引価格の最小値のファンダメンタル価格からの乖離の平均をまとめた．オーバーシュートしていない（価格がファンダメンタル価格を下回らない）場合は乖離

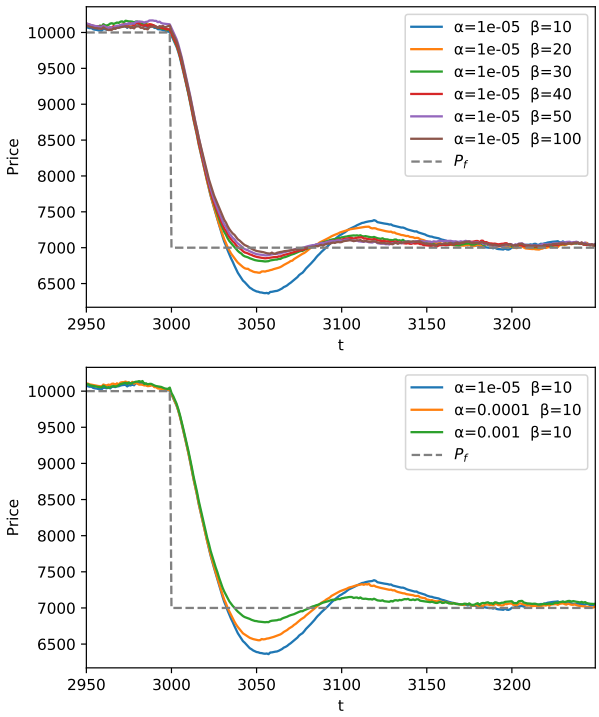


図 2: ファンダメンタル価格急落時の価格推移（上：α を固定　下：β を固定）

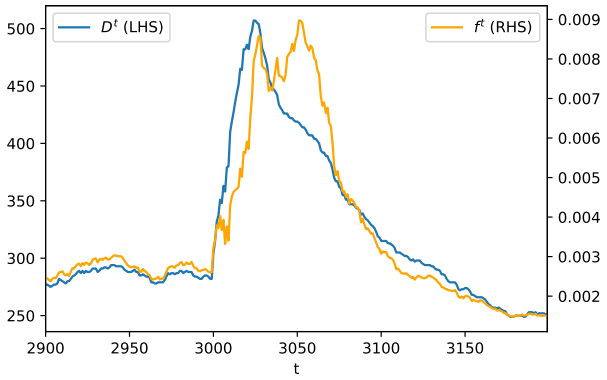


図 3: ファンダメンタル価格急落時の需要量 $D^{t'}$ と貸株料 $f^{t'}$ （$\alpha = 0.0001, \beta = 50$ の場合．$t = 3,000$ で $P_f^t = 10,000$ から 7,000 に下落）

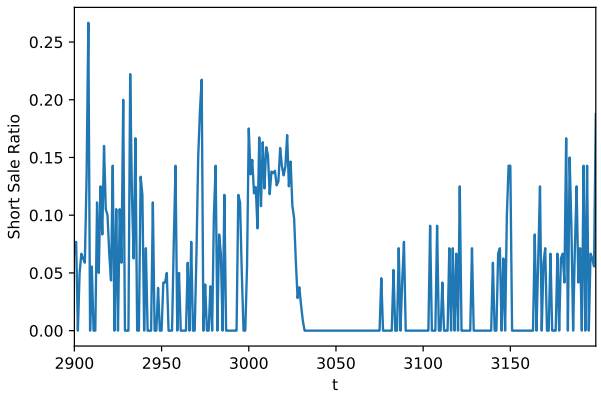


図 4: 各時点での全約定回数に占める空売りの比率（$\alpha = 0.0001, \beta = 50$ の場合．$t = 3,000$ で $P_f^t = 10,000$ から 7,000 に下落）

は 0 とした．この結果から，貸借市場の流動性を表すパラメータ α, β が大きくなるほど（流動性が低くなるほど），ファンダメンタル価格急落時のオーバーシュートは小さくなることが示せた．

表 3: オーバーシュートの大きさ

		β					
		10	20	30	40	50	100
α	1e-05	771.88	462.84	313.48	243.28	200.86	139.28
	0.0001	572.32	316.62	219.14	175.42	156.62	125.48
	0.001	288.52	174.96	143.82	131.38	122.22	109.78

5　まとめ

本研究では株式の売買市場に加え，上記の流動性の観点を貸借市場に加えた連成型人工市場を構築し，貸借市場の流動性の変化，特に貸株料の変化が売買市場の効率性に与える影響を調べた．

ファンダメンタル価格一定（平常時）と急落時の 2 つの条件でシミュレーションを行うことで，貸借市場の流動性が高い場合には，平常時には取引価格のファンダメンタル価格からの乖離が小さいという意味で売買市場は効率性が高まるが，ファンダメンタル価格急落時には自由な空売りが取引価格をオーバーシュートさせてしまうことが分かった．一方，貸借市場の流動性が低い場合には，ファンダメンタル価格急落時に空売りへの制約が自然に強まることで取引価格のオーバーシュートが小さくなるという意味で，空売り規制と同様に市場の効率性を高める一方，平常時には取引価格のファンダメンタル価格からの乖離を大きくしてしまうことが分かった．

本研究の寄与としては，金融市場の制度設計の際などに活用される人工市場に，需給に基づく貸株のメカニズムを導入した点にある．当モデルを用いたシミュレーションからは，貸借市場の流動性による制約は空売り規制と同様の効果をもたらすことが示された．これは，貸借市場の流動性が十分に高い場合には問題にならないだろうが，貸借市場の流動性が低い銘柄においても同様の空売り規制を設けてしまうと空売りに対して二重の制約を課すことになり，予期せず市場の効率性を損なってしまう可能性があることを示している．本研究で提示した貸借市場を導入した人工市場により，貸借市場の流動性を考慮した空売り規制のシミュレーションが可能となり，これは今後の課題として考えられる．また，本研究では一定とした各エージェントの貸株比率の変化の検証についても今後の課題となる．貸株は性質上すぐに売る可能性のある株式に対して行うことは難しいためアクティブファンドで行うことは難しく [10]，貸株を行う主体の多くはアクティブではなくパッシブ運用により運用されている株式だと考えられ，パッシブ運用の比率は市場全体の貸株比率に影響すると考えられる．例えば日本銀行によるパッシブ運用を行う ETF の買い入れは市場の効率性を損なうのではないかという議論があるが，[11] で指摘されているように，パッシブ運用からの貸株の流動性供給は売買市場の効率性向上に貢献する可能性があり，本研究で提案したモデルではそのようなパッシブ運用比率の上昇に伴う貸株比率の変化まで考慮した市場の効率性の変化を検証することができ，今後の課題として考えられる．

参考文献

[1] 八木勲, 水田孝信, 和泉潔．人工市場を用いた株式市場における空売り規制の有効性分析」．第 4 回 人工知能学会ファイナンスにおける人工知能応用研究会（SIG-FIN）SIG-FIN-004-02, pp. 7–14, 2010.

[2] 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 C (電子・情報・システム部門誌), Vol. 133, No. 9, pp. 1694–1700, 2013.

[3] GeneD ’avolio. The market for borrowing stock. *Journal of financial economics*, Vol. 66, No. 2-3, pp. 271–306, 2002.

[4] 宇野淳, 梅野淳也, 室井理沙．日本株レンディング市場の実証分析–株券貸借モデルによる空売り規制効果の測定 (特集 市場流動性)．証券アナリストジャーナル, Vol. 47, No. 6, pp. 19–33, 2009.

[5] Pedro AC Saffi and Kari Sigurdsson. Price efficiency and short selling. *The Review of Financial Studies*, Vol. 24, No. 3, pp. 821–852, 2010.

[6] 川久保佐記, 和泉潔. 連成型人工市場によるオプション取引戦略が原資産市場に与える影響の分析. 人工知能学会論文誌, Vol. 31, No. 6, pp. AG–D, 2016.

[7] 水田孝信. 株式と etf の裁定取引にかかるコストと流動性の関係—人工市場によるシミュレーション分析—. 人工知能学会全国大会論文集 一般社団法人 人工知能学会, pp. 2O1J1302–2O1J1302. 一般社団法人 人工知能学会, 2019.

[8] 大井朋子. エージェントシミュレーションを用いた 「価格規制」 と「ネイキッド・ショート・セリングの禁止」 の有効性の検証. FSA リサーチレビュー, Vol. 7, , 2013.

[9] Carl Chiarella, Giulia Iori, and Josep Perell. The impact of heterogeneous trading rules on the limit order book and order flows. *Journal of Economic Dynamics and Control*, Vol. 33, pp. 525–537, 03 2009.

[10] 水田孝信. 信託報酬ゼロの出現～コスト以上に重要なこと（スパークス・アセット・マネジメント株式会社　スペシャル･レポート　2018/11/07）. https://www.sparx.co.jp/report/special/2411.html.

[11] 明田雅昭ほか. パッシブ運用のエンゲージメント一論点整理と提案一. CGSA フォーラム, Vol. 16, pp. 13–32, 2018.

人工市場を用いたレバレッジドETFがザラ場市場に与える影響分析

Analysis of the impact of leveraged ETF in continuous double auction market using an agent-based simulation

丸山隼矢[1*] 水田孝信[2] 八木勲[3]
Maruyama Shunya[1] Mizuta Takanobu[2] Yagi Isao[3]

[1] 神奈川工科大学大学院工学研究科情報工学専攻
[1] Course of Information and Computer Sciences, Graduate School of Kanagawa Institute of Technology
[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.
[3] 神奈川工科大学情報学部情報工学科
[3] Faculty of Information Technology, Kanagawa Institute of Technology

Abstract: レバレッジドETFとは，日々のリターンが原資産や原指数（例えば，日経平均やTOPIXなど）の価格の変動率に一定の倍数を乗じた値動きをするETFのことを指す．レバレッジドETFは，レバレッジ率を維持する（保有する原資産の純資産総額を，レバレッジドETFの純資産総額の決められた倍数に維持する）よう，原資産の価格が上昇すれば原資産を買い，反対に下落した際は原資産を売るというリバランス取引を日々行わなければならない．そのため，これらの売買が原因で原資産の価格を不安定にさせているのではないかと言われている．これまでに，人工市場を用いた研究によってレバレッジドETFが板寄せ方式の原資産市場の価格形成に影響を与えることが知られているが，ザラ場方式の市場は未調査のままである．そこで本研究では，ザラ場方式の原資産市場においてレバレッジドETFが価格形成に与える影響を調査した．その結果，レバレッジドETFのリバランス取引の最低注文数が小さいほど市場の価格形成に与える影響が大きいことを確認した．

1 まえがき

レバレッジドETFとは，日々のリターンが原資産や原指数（例えば，日経平均やTOPIXなど）の価格の変動率に一定の倍数を乗じた値動きをするETFを指す．これらのETFは，金融市場指数に投資するため個別株式銘柄よりもローリスクでありながら，レバレッジの恩恵により比較的高いリターンが望めるため，運用残高が飛躍的に伸びてきている．

しかし，レバレッジドETFは，あらかじめ設定されたレバレッジ率を維持する（保有する原資産の純資産総額を，レバレッジドETFの純資産総額の決められた倍数に維持する）よう取引を行わなければならない．この取引のことをリバランス取引という．

リバランス取引について具体例を挙げて説明する．ある資産の2倍に乗じた動きをするレバレッジドETFを考える．今原資産価格が100とし，ETFの元手が100のキャッシュのみとすると，ETFの初期純資産総額は先物などを原資産として保有することにより200となる（不足する元手100は借金することになる）．ここで，原資産価格が100から110に10％上昇したとすると，ETFの純資産総額が20％上昇するように原資産を保有しなければならない．つまり，純資産総額は240必要となる．しかし，原資産価格の上昇は10％なので，ETFの実際の純資産総額は220にしかならず，20少ない状態となる．そのため，原資産を20追加購入しなければならない．反対に，原資産価格が100から90に10％下降したとすると，ETFの純資産総額が20％下降するように原資産を保有しなければならない．つまり，純資産総額は160となる．しかし，原資産価格の下降は10％なので，ETFの実際の純資産総額は180にしかならず，20多い状態となる．そのため，原資産を20追加売却しなければならない．

このように，レバレッジ率を維持するため日々原資産の価格が上昇すれば原資産を買い，反対に下落した際は原資産を売るため，リバランス取引が原資産の価格を不安定にさせているのではないかと言われている[Childs 14].

*連絡先：神奈川工科大学大学院工学研究科情報工学専攻
〒243-0292 神奈川県厚木市下荻野1030
E-mail: s1885023@cce.kanagawa-it.ac.jp

レバレッジド ETF が原資産に与える影響については，これまでにさまざまな実証研究が行われてきた [Cheng 09, Schwert 90, Deshpande 09, Trainor Jr. 10]．例えば，Rompotis[Rompotis 17] は，英国市場においてレバレッジド ETF はレバレッジの維持はできるもののトラッキングエラー (ポートフォリオのリターンとベンチマークとの乖離) が発生し，トラッキングエラーが市場のボラティリティの上昇と関連していることを示唆している．また，Ben-David et al.[Ben-David 16] は，レバレッジド ETF がリバランス取引を円滑に行うためには，大口の投資家（高頻度取引，ヘッジファンド，および指定参加者）の存在が不可欠とし，市場混乱時にそうした大口の投資家とレバレッジド ETF の取引により市場の混乱が悪化することを示唆している．

このように，レバレッジド ETF が市場に与える影響を調査したものは多く存在する．しかし，金融市場の形成にはさまざまな事象と要因が複雑に関わっているため，これらの結果がレバレッジド ETF 固有の影響なのか判断することは難しい．

特定の取引が市場に与える影響を分析するような課題，つまり，実証研究では議論が困難な課題を分析する方法の 1 つとして，人工市場を用いる方法がある．人工市場とは，コンピュータ上に仮想的に構築されたマルチエージェントシステムの金融市場である [Chiarella 09, Chen 12, Yeh 13]．人工市場では，エージェントにそれぞれ独自の売買手法を与え，それらを投資家として金融資産の取引をさせると，市場がどのような振る舞いをするかを確認することができる．

最近では人工市場を用いて市場規制制度が金融市場に与える本質的な影響を分析する研究が盛んに行われており [Yagi 10, Yeh 13, Mizuta 15, Zhou 17]，レバレッジド ETF に関しても次のような研究が行われている．Yagi et al.[Yagi 16] は，価格決定メカニズムが板寄せ方式の原資産市場においてレバレッジド ETF が原資産市場の価格形成に与える影響について検証している．その結果，レバレッジド ETF の運用資産が大きくなるほど，リバランス取引量が増し，ボラティリティが増加することを確認している．

しかし，価格決定方式がザラ場方式の原資産市場においてレバレッジド ETF 運用者の取引手法や原資産市場の価格形成に着目した議論はされていない．リバランス取引を一度に行うか日中を通して分散して行うかは運用者の裁量に任されている．リバランス取引を一度に行うと，運用資産の大きなレバレッジ ETF ほど市場の価格形成に与える影響が大きくなるため [Yagi 16]，リバランス取引は複数回に分けて行う方がよいかもしれない．

よって本研究では，ザラ場方式の原資産市場においてリバランス取引を複数回に分けて行うとき市場の価格形成にどのような影響を与えるか調査する．リバランス取引は，市場価格の変化に伴い変化するレバレッジド ETF のレバレッジ率と，予め設定されたレバレッジ率の乖離の大きさが一定の値を超えたとき行う．本研究におけるリバランス取引は，そうした乖離した時の注文数がある閾値を超えたときのみ発注するようにした．そして，この閾値（以降，注文数閾値と呼ぶ）を変化させ実験を行う．

その結果，注文数閾値が小さい場合，ボラティリティや市場非効率性が増加することが確認できた．一方で，注文数閾値が大きい場合は，運用資産規模が変化してもボラティリティに影響を与えないことが確認できた．

2 人工市場モデル

本研究では，水田ら [Mizuta 15] の人工市場モデルを基にして，新たにレバレッジを維持するためのリバランス取引を行うエージェント（以下，レバレッジド ETF エージェントと記す）を追加した．

2.1 注文プロセス

価格決定メカニズムは連続ダブルオークション方式（ザラ場方式）である．本市場に参加するエージェントは，通常エージェントとレバレッジド ETF エージェントの 2 種類である．通常エージェントが注文後，レバレッジド ETF エージェントがリバランス取引を行うか判定し，必要であれば注文を行う．時刻 t は通常エージェントが注文を出す，又はレバレッジド ETF エージェントが注文を出すごとに 1 増える．これを t_{max} まで繰り返し行う．

2.2 通常エージェント

本市場には，n 体のエージェントが取引に参加する．

エージェント j は注文価格，売り買いの別を以下のように決める．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）$r_{e_j}^t$ は，以下の式（1）から得られる．

$$r_{e_j}^t = \frac{w_{1_j}^t r_{e1,j}^t + w_{2_j}^t r_{e2,j}^t + u_j \epsilon_j^t}{w_{1_j}^t + w_{2_j}^t + u_j} \tag{1}$$

ここで，$w_{i_j}^t$ は時刻 t，エージェント j の重みであり，シミュレーション開始時に，それぞれ 0 から $w_{i,max}$ までの一様乱数で決める．この値は後述する学習過程により変化する．u_j はエージェント j の重みであり，シミュレーション開始時に，それぞれ 0 から u_{max} までの一様乱数で決められるが，シミュレーション中に変化することはない．

${r_{e1,j}}^t$ は，下記に示す．時刻 t におけるエージェント j のファンダメンタル投資家成分だけの予想リターンで ${r_{e1,j}}^t = \log(P_f/P^{t-1})$，同様に ${r_{e2,j}}^t$ はテクニカル投資家成分だけの予想リターンで，${r_{e2,j}}^t = \log(P^{t-1}/P^{t-1-\tau_j})$ である．P_f は時間によらず一定のファンダメンタル価格である．ファンダメンタル価格とは，金融資産を発行する企業自体が持っている実態の価値に基づいた価格を指す．P^t は時刻 t での取引価格（取引されなかった時刻では一番最近に取引された価格であり，時刻 $t=0$ では $P^t = P_f$ とする）である．τ_j は 1 から τ_{max} までの一様乱数でエージェントごとに決める．ϵ_j^t は時刻 t，エージェント j の乱数項で，平均 0，標準偏差 σ_e の正規乱数である．

式（1）の右辺の分子の第 1 項目は，ファンダメンタル価格と比較して，安ければプラスの予想リターンを，高ければマイナスの予想リターンを示すファンダメンタル投資家の成分である．なお，ファンダメンタル投資家は通常ファンダメンタル価値を参照して投資判断を行う投資家を指す．第 2 項目は過去のリターンがプラス（マイナス）であるならプラス（マイナス）の予想リターンを示す，テクニカル投資家の成分である．テクニカル投資家は過去の価格推移を参照して投資判断を行う投資家を指す．第 3 項目はノイズを表している．

予想価格 ${P_{e_j}}^t$ は，予想リターン ${r_{e_j}}^t$ に基づいて式（2）から求められる[1]．

$$ {P_{e_j}}^t = P^{t-1} \exp({r_{e_j}}^t) \tag{2} $$

注文価格 ${P_{o_j}}^t$ は ${P_{e_j}}^t - P_d$ から ${P_{e_j}}^t + P_d$ で決定される．ここで，P_d は平均 0，標準偏差 σ_p の正規乱数に予想価格をかけたものである．価格の変化幅の最小単位は δP とし，注文の売り買いは予想価格 ${P_{e_j}}^t$ と注文価格 ${P_{o_j}}^t$ の大小関係で式（3）のように決まる[2]．

$$ \begin{aligned} &{P_{e_j}}^t > {P_{o_j}}^t \text{なら買い} \\ &{P_{e_j}}^t < {P_{o_j}}^t \text{なら売り} \end{aligned} \tag{3} $$

注文数は常に 1 と一定とする．

エージェントは状況に応じて戦略を切り替える学習を行う．学習プロセスは以下のようにモデル化した．価格変化の方向を当てている戦略のウエイトを引き上げ，外れている戦略のウエイトを引き下げるようにした．また，学習期間のリターン ${r_l}^t$ をかけることにより，小さい価格変動を当てたり外したりしても大きくウエイトが増減しないようにした．学習はエージェントごとに，各エージェントが注文を出す直前に行われる．これらの予想リターン ${r_{e_i,j}}^t$ が学習期間のリターン ${r_l}^t = \log(P^{t-1}/P^{t-t_l})$ と比べ，同符号なら，${w_i}_j^t$ を ${w_i}_j^{t-1} + k_l|{r_l}^t|q_j^t(w_{i,max} - {w_i}_j^{t-1})$ に，異符号なら，${w_i}_j^t$ を ${w_i}_j^{t-1} - k_l|{r_l}^t|q_j^t {w_i}_j^{t-1}$ とする．ここで，k_l は定数，q_j^t は時刻 t，エージェント j に与えられる 0 から 1 までの一様乱数である．

上記学習プロセスの他に，小さい確率 m で ${w_i}_j^t$ を再設定する．つまり，0 から $w_{i,max}$ までの一様乱数で決めなおす．これはランダム学習を意味しており，実績からの学習と組み合わせることにより，エージェントが試行錯誤的により良い戦略のウエイトを求める姿を客観的にモデル化している．

2.3 レバレッジド ETF エージェント

原資産価格が上昇もしくは下落するとき，レバレッジド ETF のレバレッジ率が変化する．それに伴いリバランス取引で発注する注文数も変動し，その注文数が注文数閾値を超えたときレバレッジド ETF エージェントは注文を行う．

第 t 期のレバレッジド ETF エージェントの保有純資産額 NAV^t は次のように表される．

$$ NAV^t = P^{t-1}S^t + C^t \tag{4} $$

なお，初期資産保有数 S^0 は，初期原資産価格 P^0 と維持すべきレバレッジ率 L を基に求められる．$S^0 = LC^0/P^0$ とし，（分子が初期純資産総額に該当する），初期キャッシュC^0 の値を $C^0 = C^0 - P^0S^0$ のように更新する（初期キャッシュのレバレッジ L 倍で資産運用するため，キャッシュの $(L-1)$ 倍分の原資産は借金して購入することを意味している）．第 t 期のレバレッジド ETF エージェントのリバランス取引に必要な注文数 V^t は次のように表される．ただし，注文数に端数が出る場合は切り捨てる．

$$ V^t = \lfloor \frac{(L-1)\,P^{t-1}S^t + LC^t}{P^{t-1}} \rfloor \tag{5} $$

$C^t \cdot S^t$ は，それぞれ第 t 期のレバレッジド ETF エージェントのキャッシュと資産保有数を表す．上記の注文数の絶対値が，閾値 V_{thr} を超えたとき注文を行う．V^t が負の場合は売り注文を行い，正の場合は買い注文を行う．また，本エージェントの注文は，レバレッジ率を確実に維持するために成り行き注文で行う．本研究では，こうしたレバレッジド ETF エージェントが 1 体参加する．

[1] 本研究では対数リターンを使用している．そのため予想リターンは現在の価格の対数と予想価格の対数の差である．すなわち，${r_{e_j}}^t = \log {P_{e_j}}^t - \log P^{t-1} = \log {P_{e_j}}^t/P^{t-1}$ であり，これより式（2）が導き出される．

[2] 予想価格の高低に関わらず，買い注文と売り注文が注文される確率は同じであるが，取引が成立する確率は必ずしも同じではない．一般に予想価格が高いエージェントの注文は，買い注文の方が売り注文より取引が成立しやすく，反対に予想価格が低いエージェントの注文は，売り注文の方が買い注文より取引が成立しやすい．

表 1: 各条件下における注文数閾値 V_{thr}

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	1	2	3	4	5
	20	2	4	6	8	10
	30	3	6	9	12	15
	40	4	8	12	16	20
	50	5	10	15	20	25
	60	6	12	18	24	30
	70	7	14	21	28	35
	80	8	16	24	32	40
	90	9	18	27	36	45
	100	10	20	30	40	50

表 2: $C_{mag} = 10$ のときのスタイライズド・ファクト

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
尖度		5.72	4.35	2.33	1.35	1.12
	ラグ					
リターンの	1	0.283	0.242	0.200	0.172	0.165
2 乗の	2	0.227	0.183	0.134	0.112	0.105
自己相関	3	0.188	0.141	0.098	0.081	0.073
	4	0.165	0.115	0.083	0.066	0.058
	5	0.138	0.104	0.072	0.052	0.045

3 シミュレーション

実験は以下のパラメータのように設定した．$n = 1{,}000$，$P_f = 10{,}000$，$w_{1,max} = 1$，$w_{2,max} = 5$，$u_{max} = 1{,}\tau_{max} = 10{,}000$，$\sigma_e = 0.03$，$\sigma_p = 0.05$，$t_c = 15{,}000$，$t_l = 10{,}000$，$k_l = 4$，$m = 0.01$，$\delta P = 1$，$t_{max}$=1,000,000，$L = 2$.

本研究では，市場に与える影響について調査する．一つ目にレバレッジド ETF のリバランス取引を行う際の注文数閾値を変化させることによる影響を調査する．二つ目に，リバランス取引の注文数はレバレッジド ETF の純資産総額（運用資産高）の大きさからの影響も受けるため，運用資産を変化させた実験も行う．

レバレッジド ETF エージェントの運用資産は以下のように変化させる．初期キャッシュ$C^0 = 1{,}000{,}000 \times C_{mag}$ とし，初期キャッシュの倍率 C_{mag} の値を大きくすることでレバレッジド ETF エージェトの運用資産の大小を表す．C_{mag} の値は，10，20，30，40，50，60，70，80，90，100 の 10 パターンとする．

注文数閾値 V_{thr} は C_{mag} が 10 のとき，1，2，3，4，5 の 5 パターンとし，これを基準に C_{mag} を変化させたときのそれぞれの注文数閾値を決めていく．つまり，注文数閾値を初期キャッシュ量に比例して大きくする[3]．各パターンにおける注文数閾値を表 1 に示す．

上記の表を基に議論する際，便宜上 V_{nor} を用いる．V_{nor} は，注文数閾値 V_{thr} を初期キャッシュの倍率 C_{mag} で規格化した値となる[4]．これにより V_{nor} の値が同じであれば，C_{mag} が大きくなっても V_{thr} も大きくなるので，初期キャッシュ量の大小がリバランス取引に与える影響を考慮せずにすむ．

以上から V_{nor} を 0.1，0.2，0.3，0.4，0.5 の 5 パターン，C_{mag} を 10，20，30，40，50，60，70，80，90，100 の 10 パターンとし，それぞれを組み合わせシミュレーションを行う．シミュレーションは各条件の下でそれぞれ 50 回ずつ試行し，その結果をもとに議論を行った．

3.1 人工市場モデルの妥当性

現実の市場ではファット・テイルとボラティリティ・クラスタリングが現れる，と多くの実証研究で指摘されている [Sewell 06, Cont 01]．ファット・テイルは，資産価格のリターンの分布が厚い，すなわち，尖度が正であることである．また，ボラティリティ・クラスタリングは資産価格のリターンの 2 乗が，ラグが増えても自己相関係数が有意に正であり，長期記憶性を持つことである．

本提案モデルにおいてもファット・テイルとボラティリティ・クラスタリングが再現されているか確認を行った[5]．その結果，すべての条件下においてこれらが再現されることが確認できた．一例として，C_{mag} が 10 のとき，かつ，V_{thr} の全てのパターンのときのスタイライズドファクトを表 2 に記す．この表からもわかるように，尖度とリターンの 2 乗の自己相関ともに正値となっている．なお, スタイライズド・ファクトはすべて 100 ステップ間隔での価格を用いて計算した．なぜならば，本モデルの時刻は注文をしただけで取引が成立しない場合も時刻が進むため，時刻 1 ごとの全ての価格を用いたスタイライズド・ファクトでは，多くの価格変動がないデータによるバイアスがかかってしまうからである．

[3] このような対応をとった理由は次のとおりである．リバランス取引の注文数は，市場価格の変化量（値幅）が同じとき，運用資産高，つまり初期キャッシュ量に比例して大きくなる．これは，初期キャッシュ量が大きいほどリバランス取引が市場に与える影響が大きくなることを意味している．よって，注文数閾値の変化させたときのリバランス取引が市場に与える影響を確認する際，初期キャッシュ量の大小による影響が反映されることを防ぐためである．

[4] V_{nor} と V_{thr} の関係は次の式で表される．$V_{thr} = V_{nor} \times C_{mag}$

[5] リバランス取引法の違いが市場に与える影響を調べることが目的であるため，レバレッジド ETF が参加していない環境の下でも再現の確認を行った．

表 3: ボラティリティ

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	0.139%	0.102%	0.089%	0.087%	0.086%
	20	0.172%	0.107%	0.090%	0.086%	0.086%
	30		0.111%	0.090%	0.086%	0.086%
	40		0.111%	0.087%	0.085%	0.086%
	50		0.113%	0.087%	0.084%	0.085%
	60			0.086%	0.084%	0.085%
	70			0.084%	0.083%	0.084%
	80			0.084%	0.083%	0.083%
	90			0.083%	0.081%	0.084%
	100			0.082%	0.081%	0.083%

表 4: 市場非効率性

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	0.243%	0.177%	0.148%	0.136%	0.130%
	20	0.389%	0.228%	0.169%	0.144%	0.137%
	30		0.274%	0.187%	0.150%	0.139%
	40		0.318%	0.198%	0.156%	0.142%
	50		0.362%	0.214%	0.161%	0.145%
	60			0.228%	0.169%	0.150%
	70			0.241%	0.174%	0.152%
	80			0.259%	0.183%	0.157%
	90			0.269%	0.189%	0.162%
	100			0.287%	0.195%	0.164%

表 5: レバレッジド ETF の注文回数

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	9879	3084	952	312	110
	20	7315	2508	847	268	99
	30		2120	740	243	95
	40		1855	638	223	89
	50		1675	587	204	85
	60			549	197	85
	70			509	185	81
	80			485	186	78
	90			468	179	78
	100			449	173	76

表 6: レバレッジド ETF の一回当たりの注文数

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	1	2	3	4	5
	20	4	5	7	8	10
	30		9	10	12	15
	40		13	14	16	20
	50		16	18	21	25
	60			22	25	30
	70			26	29	35
	80			29	34	40
	90			33	38	45
	100			36	42	50

3.2 シミュレーション結果

各パターンの市場のボラティリティを表 3 に示す[6]．表 3 より，V_{nor} が大きくなるにつれボラティリティが減少傾向にあることが確認できる．また，C_{mag} が大きくなるにつれてもボラティリティが減少傾向にあることが確認できる．しかし，例外として注文数閾値が小さい（$V_{nor} = 0.1, 0.2$）ときは，C_{mag} が大きくなるにつれボラティリティが増加傾向にあることが確認できる．

次に市場の非効率性を表 4 に示す．市場非効率性は，次式にて求められる．

$$M_{ie} = \frac{1}{t_{max}} \sum_{t=1}^{t_{max}} \frac{|P^t - P_f|}{P_f} \tag{6}$$

M_{ie} が 0 なら効率的で，大きくなるほど非効率であることを示す．

表 4 より，V_{nor} が大きくなるにつれ，市場非効率性が減少傾向にあることが確認できる．一方で，C_{mag} が大きくなるにつれ市場非効率性が増加傾向にあることが確認できた．

[6]数値未記入部分は，取引価格が一方向に偏り市場崩壊を起こしたため，値の取得が不可能であった．

4 考察

まず初めにボラティリティについて議論する．ボラティリティの変動はレバレッジド ETF の注文回数と大きな関係があると考えられる．レバレッジド ETF の注文回数を表 5 に示す．

表 5 より，V_{nor} が大きくなるにつれ注文回数が急激に減少していることが確認できる．この結果から V_{nor} が大きいと，一回の取引が市場に与える影響は大きいが，それ以上に注文回数が大きく減少したため，ボラティリティは小さくなったと考えられる．C_{mag} に関しても同様の理由でボラティリティが小さくなったと考えられる．

しかし，注文数閾値が小さい（$V_{nor} = 0.1, 0.2$）ときは例外で，注文回数は減少しているがボラティリティは増加傾向にある．これは，レバレッジ ETF の一回当たりの注文数が関係していると考えられる．レバレッジド ETF の一回当たりの注文数を表 6 に示す．表 6 より，注文数閾値が小さい（$V_{nor} = 0.1, 0.2$）ときは，C_{mag} が大きくなると，注文数閾値（表 1）より大きく注文数が増加している．この一回当たりの注文数の増加が大きいのに対し注文回数の減少幅が小さいため，ボラティリティは大きくなったと考えられる．

次に市場非効率性について議論する．市場非効率性は，レバレッジド ETF の注文回数と一回当たりの注文

表 7: スプレッド

		V_{nor}				
		0.1	0.2	0.3	0.4	0.5
C_{mag}	10	12	11	11	11	11
	20	11	10	11	11	11
	30		10	11	11	11
	40		10	10	11	11
	50		9	10	11	11
	60			10	11	11
	70			10	11	11
	80			10	11	11
	90			10	11	11
	100			10	11	11

表 8: 注文数閾値に対するリバランス取引発動までの市場価格の変化量（値幅）

	V_{nor}				
	0.1	0.2	0.3	0.4	0.5
値幅	5	10	15	20	25

数が関係していると考えられる．

表 5 より，V_{nor} が大きくなるにつれ注文回数が減少している．レバレッジド ETF のリバランス取引の注文による影響が弱まるため市場が効率的になったと考えられる．一方で，表 6 より，C_{mag} が大きくなるにつれ一回当たりの注文数が増加している．C_{mag} が大きくなるにつれ注文回数は減少するものの，それに対し一回当たりの注文数が増加するため，価格が大きく変動し，ファンダメンタル価格が収束するのに時間を要するので，市場が非効率的になったと考えられる．

注文数閾値が小さい（$V_{nor} = 0.1, 0.2$）とき市場崩壊を起こすことがあるが，これはスプレッドと関係していると思われる．スプレッドとは，これまでに発注された注文のうち取引が成立せずに市場に残された注文のうち，最も高い買い注文の価格（最良買い気配値）と最も安い売り注文の価格（最良売り気配値）の差を意味する．表 7 は市場の平均スプレッドである．市場崩壊を起こさないときは，多少の大きさの差異はあるものの安定した値となっている．表 8 は注文数閾値に対するリバランス取引発動までの市場価格の変化量（値幅）である．

表 7 と表 8 から，市場崩壊するときはリバランス取引発動までの値幅がスプレッドの値とほぼ同じかそれ以下であることがわかる．この関係性が市場崩壊の起因となっていると考えられる．その理由を具体例（$V_{nor} = 0.1$ のとき，つまり，表 8 から市場価格が ± 5 だけ変化したとき，リバランス取引が行われるとき）を挙げて説明する．レバレッジド ETF が買いのリバランス取引を行い市場価格が 10,000 となり，このときの最良買い気配値が 10,000，最良売り気配値が 10,010 である（つまり，スプレッドが 10 である）とする．ここで次にリバランス取引が行われる状況を売りと買いに分けて確認する．次の買いのリバランス取引はここからさらに価格が 5 下がる（9,995）と行われる，この状況になるには通常エージェントの多くの買い注文が約定しなければならない．一方，売りのリバランス取引は価格が 5 上がる（10,005）と行われるが，この状況は通常エージェントの売り注文が 1 つ約定すると成り立ってしまう．ここで言えることは，直近の（リバランス）取引の後，通常エージェントが反対売買注文を出しそれが約定（取引成立）すると，リバランス取引が立て続けに起こってしまうということである．この現象は，C_{mag} が大きくなるほど一回あたりの注文数も大きくなるため，発生しやすくなる．

それとは反対に注文数閾値が大きくなる（$V_{nor} = 0.3$ 以上）と，注文数閾値に対するリバランス取引発動までの市場価格の変化量はスプレッドの値より大きいため，上記のように通常エージェントの 1 回の反対売買によってリバランス取引が行われることはほぼ起こりえない．

以上より，注文数閾値が小さく C_{mag} が大きいときほど，つまり，リバランス取引発動までの値幅がスプレッドの値より小さくなるほど，リバランス取引が立て続けに行われる可能性が高くなり，市場崩壊へとつながるものと考えられる．

5 まとめ

本研究では，注文数閾値という概念を導入して，ザラ場方式の原資産市場において一回のリバランス取引で必要な注文を複数回に分けて注文するとき，市場の価格形成にどのような影響を与えるか調査した．その結果リバランス取引の注文数閾値が大きいとき，つまり最低注文数が大きいとき，ボラティリティに与える影響が小さいことを確認した．反対に，注文数閾値が小さいとき，一度に多くの注文が出され，市場価格が急激に動くためボラティリティが増加すること確認した．また，市場非効率性の面では，注文数閾値が大きくなるにつれ市場が効率的になるが，レバレッジド ETF の運用資産が大きくなると市場が非効率的になることを確認した．

そして，レバレッジド ETF の注文数閾値が小さいかつ運用資産が大きい場合は，ボラティリティが増加し，市場が非効率的になり市場崩壊を起こす可能性があることが確認できた．今回の実験により，この現象はリ

バランス取引発動までの値幅がスプレッドの値以下のとき顕著に出ることが示唆された.

参考文献

[Ben-David 16] Ben-David, I., Franzoni, F., and Moussawi, R.: Exchange Traded Funds (ETFs), Technical report, National Bureau of Economic Research (2016)

[Chen 12] Chen, S.-H., Chang, C.-L., and Du, Y.-R.: Agent-based Economic Models and Econometrics, *Knowledge Engineering Review*, Vol. 27, No. 2, pp. 187–219 (2012)

[Cheng 09] Cheng, M. and Madhavan, A.: The Dynamics of Leveraged and Inverse Exchange-Traded Funds, *Journal of Investment Management*, Vol. 7, No. 4, pp. 43–62 (2009)

[Chiarella 09] Chiarella, C., Iori, G., and Perelló, J.: The Impact of Heterogeneous Trading Rules on the Limit Order Flows, *Journal of Economic Dynamics and Control*, Vol. 33, No. 3, pp. 525–537 (2009)

[Childs 14] Childs, M.: Fink Says Leveraged ETFs May 'Blow Up' Industry (2014)

[Cont 01] Cont, R.: Empirical properties of asset returns: stylized facts and statistical issues, *Quantitative Finance*, Vol. 1, pp. 223–236 (2001)

[Deshpande 09] Deshpande, M., Mallick, D., and Bhatia, R.: Understanding Ultrashort ETFs, *Barclays Captital Special Report*, pp. 1–6 (2009)

[Mizuta 15] Mizuta, T., Izumi, K., Yagi, I., and Yoshimura, S.: Investigation of Price Variation Limits, Short Selling Regulation, and Uptick Rules and Their Optimal Design by Artificial Market Simulations, *Electronics and Communications in Japan*, Vol. 98, No. 7, pp. 13–21 (2015)

[Rompotis 17] Rompotis, G. G.: Empirical Insights on the Trading Behavior of the UK Leveraged ETFs, *Journal of Financial Innovation*, Vol. 1, No. 3, pp. 159–197 (2017)

[Schwert 90] Schwert, G. W.: Stock Volatility and the Crash of 87, *The Review of Financial Studies*, Vol. 3, No. 1, pp. 77–102 (1990)

[Sewell 06] Sewell, M.: Characterization of financial time series (2006)

[Trainor Jr. 10] Trainor Jr., W. J.: Do Leveraged ETFs Increase Volatility, *Technology and Investment*, Vol. 1, pp. 215–220 (2010)

[Yagi 10] Yagi, I., Mizuta, T., and Izumi, K.: A Study on the Effectiveness of Short-selling Regulation in View of Regulation Period Using Artificial Markets, *Evolutionary and Institutional Economics Review*, Vol. 7, No. 1, pp. 113–132 (2010)

[Yagi 16] Yagi, I. and Mizuta, T.: Analysis of the Impact of Leveraged ETF Rebalancing Trades on the Underlying Asset Market Using Artificial Market Simulation, in *12th Artificial Economics Conference*, pp. 1–11 (2016)

[Yeh 13] Yeh, C.-H. and Yang, C.-Y.: Do Price Limits Hurt the Market?, *Journal of Economic Interaction and Coordination*, Vol. 8, No. 1, pp. 125–153 (2013)

[Zhou 17] Zhou, X. and Li, H.: Buying on Margin and Short Selling in an Artificial Double Auction Market, *Computational Economics*, pp. 1–17 (2017)

DSGE によるニューラルネットの正則化と経済予想

Combining a DSGE model with Neural Networks for Forecasting Economic Variables

塩野剛志 [1]

Takashi Shiono[1]

[1] クレディ・スイス経済調査部
[1]Credit Suisse Economics Research

Abstract: This paper examines the possibility of combining a DSGE model and neural networks to supplement each other, with regard to out-of-sample forecasts for economic variables. The aim is to build a model with theoretical interpretability and state-of-the-art performance. The novel neural-net structure of TDVAE (Temporal Difference Variational Auto-Encoder) proposed by Gregor et.al [2019] enables to realize this idea. TDVAE virtually replicates a Gaussian stochastic state-space model through combination of neural networks. Because a DSGE model provides theoretical restrictions on the state transition and observation matrices of a linear state-space model, I choose to transplant those DSGE-oriented matrices into the formulations of state transition and observation probabilities in TDVAE. This TDVAE-DSGE approach certainly achieved the superior performance in the task of out-of-sample forecasts on Japan's real GDP during 1Q/2011 and 4Q/2018.

1. Introduction

This paper examines the possibility of combining a DSGE model and neural networks to supplement each other, with regard to out-of-sample forecasts for economic variables. The aim is to build a model with theoretical interpretability and state-of-the-art performance.

1.1. Forecasts by DSGE models

The dynamic stochastic general equilibrium (DSGE) model is a central tool to analyze fluctuation of macroeconomic variables in the mainstream macroeconomics, and has increasingly been used for forecasting purpose. There is a plentiful amount of literature examining the forecast performances of DSGE models. Following the seminal work by Smets and Wouters [2007], Edge and Gurkaynak [2010] reported that the forecast accuracy of medium-scale DSGE model was not satisfactory but so are the other statistical model (Bayesian VAR) and professional forecasts.

The linearized DSGE model can virtually be regarded as a linear state-space model, with the theoretical specification of transition and observation matrixes. Indeed, Herbst and Schorfheide [2012] indicated that forecasts from a DSGE model could outperform those from an unrestricted VAR if the theoretical restriction on comovements among economic variables is well consistent with the actual data generation processes. Recently, Del Negro et.al [2013] showed that considering financial friction into a DSGE could contribute to the better forecast performance during the periods just after the financial crisis in 2008-09 in the US. The latest study of Hasumi et.al [2018] also found that an appropriate model could vary with underlying economic situations, through the introduction of time-varying pooling method of multiple models. They reported that the DSGE model with the financial friction outperformed the frictionless benchmark in terms of forecast performance during the 1980s financial bubble and following burst in Japan, while the frictionless model marked better scores for more recent periods. This may suggest that any single type of DSGE seems difficult to persistently outperform more flexible type of statistical models. Having said that, a DSGE model has own advantage as a descriptive tool, with the theoretical correspondence and high interpretability.

1.2. Forecasts by neural networks

One of the prominent progress in statistical modeling

in these days is so-called artificial intelligence, more precisely, deep-neural-network type of models. Their superior performances to other competing statistical models were reported in many fields. Research applying neural networks to economic time-series are quite recent, and most of them deal with the forecasts of asset prices. The early research by Chen et al. [2003] applied a neural network to predict return of Taiwanese equity index, reporting the improved accuracy compared to a GMM approach and random walk. More recently, Fischer and Krauss [2017] compared the prediction performances of LSTM (Long Short Term Memory), Random Forests, and logistic regression, and reported the superiority of LSTM in respect of the daily S&P500 data from 1992 to 2015. Borovkova and Tsiamas [2018] employed an ensemble (dynamic pooling) method of multiple LSTM networks for the intraday stock price prediction, and reported the better performance against the logistic regression benchmarks. Those pieces of research show that neural-net type of models certainly tend to beat traditional statistical models in the context of forecasting financial data, and LSTM network particularly exhibits the superior performance. LSTM network can treat non-linear dynamics and long-memory feature of time series flexibly, which differentiate from traditional statistical models like simple Vector Auto-Regression (VAR).

Although neural network type of models tend to show state-of-the-art forecast performance, it has an disadvantage of so-called "black box" feature. For most neural networks, it is generally difficult to trace the transmission process of input information to outlay with their repeated non-linear transformations, which also hampers the models to hold theoretical correspondence.

Summing up the above literatures, the theoretical restriction by DSGE could make the linear state-space model interpretable, and contribute to the forecast accuracy as long as the restriction is consistent with the actual data generation process. Nonetheless, its absolute performance is said to not satisfactory. On the other hand, neural network models tend to outperform traditional statistical models in forecasts of financial time series, but its opacity or "black box" nature of information processing hinders the application to macroeconomic study.

1.3. State-space model constructed by neural networks

The novel structure of TDVAE (Temporal Difference Variational Auto-Encoder) proposed by Gregor et.al [2019] could help to address both issues of the opacity for neural networks and the poor accuracy for DSGE models, through combining the two approaches. The purpose of this paper involves this idea, namely, constructing theoretically interpretable and state-of-the-art performance model.

As mentioned more precisely later, TDVAE enables to replicate a Gaussian stochastic state-space model in a framework of end-to-end neural networks. Generally speaking, Variational Auto-Encoder (VAE) consists of encoder and decoder, where the former transforms data to lower-dimensional state variables (so-called representation) while the latter tries to reconstruct the data from representation, with assumption for the state variables to follow a normal distribution. TDVAE extends this concept to treat dynamic state variables by analogy with state space modeling. In this TDVAE model, a filtering probability constitutes the encoder which transforms observed data to contemporaneous states variables, while a transition probability describes dynamics of states variables. Then, a smoother probability validates the transition of state variables backward. These probabilities are parameterized by means and variances of Gaussian distribution, which are expressed as the outputs of deep neural networks trained by the data, inheriting the concept of Variational Auto-Encoder.

As seen above, TDVAE is virtually a semi-parametric non-linear state-space model, in which latent states, a low of motion, and an observation function can be specified in consistent with the theoretical restrictions derived from a DSGE model.

In the next section, I explain a basic structure of the proposed TDVAE-DSGE method. Then, results from forecasting back-tests are presented in Section 3, and the backgrounds of accuracy gains are discussed in Section4.

2. Proposed Methods

This section explains a basic structure of TDVAE, and the way to specify its components to be consistent with an orthodox DSGE model.

2.1. TDVAE structure

TDVAE assumes a general form of stochastic state-space model, where the joint likelihood of a state sequence $\mathbf{z} = (z_1, \dots, z_T)$ and an observation sequence $\mathbf{x} = (x_1, \dots, x_T)$ can be written as $p(\mathbf{x}, \mathbf{z}) = \prod_t p(z_t|z_{t-1})p(x_t|z_t)$. Here, $p(z_t|z_{t-1})$ represents a transition model and $p(x_t|z_t)$ does a decoder or an observation function.

In order to optimize the data likelihood $p(\mathbf{x})$, a state-space model is commonly trained by computing a posterior $q(\mathbf{z}|\mathbf{x})$ over the states given the observations, which can be autoregressively decomposed as $q(\mathbf{z}|\mathbf{x}) = \prod_t p(z_t|z_{t-1}, \phi_t(\mathbf{x}))$, where ϕ_t is a function augmenting the observed data up to t $(x_1, \dots, x_t)$ for filtering posteriors, or the entire sequence $\mathbf{x}$ for smoothing posteriors.

For a given t, TDVAE evaluates the conditional likelihood $p(x_t|x_{<t})$ by inferring over only two latent states: z_{t-1} and z_t, and obtains the ELBO (Evidence Lower BOund) for a state-space model as follows:

$$\begin{aligned}\log p(x_t|x_{<t}) \geq \mathbb{E}_{(z_{t-1},z_t)\sim q(z_{t-1},z_t|x_{\leq t})}[&\log p(x_t|z_t) \\ &+ \log p(z_{t-1}|x_{<t}) + \log p(z_t|z_{t-1}) \\ &- \log q(z_t|z_{\leq t}) - \log q(z_{t-1}|x_{\leq t})].\end{aligned} \tag{1}$$

Maximizing ELBO (the right-hand side) instead of the likelihood itself (the left hand side) is common approach to train encoder-decoder type of neural networks such as VAE (Variational Auto-Encoder). More formal derivation of ELBO for stochastic state-space model is explained in Buesing et al. [2018].

Here, TDVAE introduces an online belief state b_t as a summary statistics of the observed (past) data at any given t. This enables to rewrite the conditional data distribution at t as $p(x_{t+1}, \dots, x_T|x_1, \dots, x_t) \approx p(x_{t+1}, \dots, x_T|b_t)$, and the filtering posteriors at t as $q(z_t|x_1, \dots, x_t) \approx p_B(z_t|b_t)$, with $b_t = f(b_{t-1}, x_t)$. Now, the loss function of TDVAE derived from belief based ELBO is obtained as follows:

$$\begin{aligned}-\mathcal{L} = \mathbb{E}_{\substack{z_t\sim p_B(z_t|b_t) \\ z_{t-1}\sim q(z_{t-1}|b_t,b_{t-1})}}[&\log p(x_t|z_t) \\ &+ \log p_B(z_{t-1}|b_{t-1}) + \log p(z_t|z_{t-1}) \\ &- \log p_B(z_t|b_t) \\ &- \log q(z_{t-1}|z_t, b_t, b_{t-1})].\end{aligned} \tag{2}$$

For actual training of TDVAE, functional forms need to be specified. The online belief state is expressed by a standard LSTM network: $b_t = LSTM(b_{t-1}, x_t)$. The mapping operator from x to a normal distribution with mean $\mu(x)$ and log-standard deviation $\log\sigma(x)$ is denoted as D, where $[\mu, \log\sigma] = \mathrm{W}_3 \tanh(W_1 x + B_1)\,\mathrm{sigmoid}(W_2 x + B_2) + B_3$, with W_1, W_2, W_3 as weight matrices and B_1, B_2, B_3 as biases. The size of hidden layer of the D maps is set to 100, and belief states have size 100. Then, I choose to use the Adam optimizer.

The set of equations describing the system and its loss function over the pair of two time periods t and $t-1$ are as follows:

$$\begin{aligned}
&b_t = LSTM(b_{t-1}, x_t) \\
&z_t^B \sim p_B^t = D(b_t) \\
&p_B^{t-1} = D(b_{t-1}) \\
&z_{t-1}^S \sim q_S^{t-1|t} = D(b_t, z_t^B) \\
&z_t^T \sim p_T^{t|t-1} = D(z_{t-1}^S) \\
&p_D = D(z_t^B) \\
&L_{t-1} = KL\left(q_S^{t-1|t} | p_B^{t-1}\right) \\
&L_t = \log p_B^t(z_t^B) - \log p_T^{t|t-1}(z_t^T) \\
&L_x = -\log p_D(x_t) \\
&Loss = L_{t-1} + L_t + L_x.
\end{aligned} \tag{3}$$

2.2. TDVAE-DSGE

As mentioned earlier, a standard New Keynesian DSGE model can be reduced to a linear state space representation, by solving its log-linearized equilibrium conditions via the method like Sims [2001] or Klein [2000]. In other words, a standard DSGE model can be summarized into a pair of the transition matrix on state variables and the observation (decoder) matrix to translate the states to observations. The values of those matrices are fully determined by the structural (deep) parameters of the DSGE model to be used.

The idea of this paper is to transplant those transition and decoder matrixes of a standard New Keynesian DSGE model into the aforementioned TDVAE framework. More precisely, means of a state transition probability

$p_T^{t|t-1}$ and a decoder probability p_D in TDVAE are to be calculated based on the given transition matrix F and the decoder matrix P of a reduced DSGE model. The system is described as follows:

$$
\begin{aligned}
&b_t = LSTM(b_{t-1}, x_t) \\
&z_t^B \sim p_B^t = D(b_t) \\
&p_B^{t-1} = D(b_{t-1}) \\
&z_{t-1}^S \sim q_S^{t-1|t} = D(b_t, z_t^B) \\
&\mu_{z^T} = F z_{t-1}^S \\
&\log \sigma_{z^T} = D(z_{t-1}^S) \\
&z_t^T \sim p_T^{t|t-1} = N(\mu_{z^T}, \sigma_{z^T}) \\
&\mu_x = P z_t^B \\
&\log \sigma_x = D(z_t^B) \\
&p_D = N(\mu_x, \sigma_x).
\end{aligned}
\tag{4}
$$

This can also be trained by using the same ELBO loss function. Hereafter, this approach is named as TDVAE-DSGE.

2.3. The DSGE model to be used

The DSGE model to be used is a standard New-Keynesian model of Smets and Wouters [2007], which over the past decade has been a standard tool in mainstream macroeconomic research.

This model includes seven structural (i.e. theoretically identified) shock processes, namely (1) a productivity shock, (2) a shock to household (investor) preferences (the time discount rate), (3) a shock to exogenous demand (government spending and foreign demand), (4) a shock to capital investment efficiency, (5) a price mark-up shock, (6) a wage mark-up shock, and (7) a discretionary monetary policy shock.

Given the materialized values of seven shocks and the deep parameters specifying behaviors of agents, the ten variables of interest, namely capital, labor inputs, production, consumption, investment, the real capital yield, the real capital price, the inflation rate, the real wage, and the nominal interest rate, are determined through the structural equations which represents optimizing behavior of the six types of economic agents, namely households (end investors), employment agencies, intermediate good producers, final good producers, capital producers, and the public sector (central bank and government).

By solving the structural equations into reduced forms, the DSGE is expressed in state-space representation, and its deep parameters, endogenous variables, and shock processes are simultaneously estimated by a Particle Filter algorithm so as to obtain the best replication of actual fluctuations in macroeconomic data. This paper denotes this standard estimation procedure as PF-DSGE for the later use. Obtaining filtered time series of these structural shock processes will enable us to identify the exogenous causes of macroeconomic fluctuations.

One major distinction of TDVAE-DSGE from PF-DSGE is an existence of belief state, which enables to compress a large number of observable data in efficient way through the LSTM network and conditions the filtering probability of state variables (i.e. shock processes). There are some pieces of research tackling data-rich estimation of a DSGE model through Dynamic Factor Model (Iiboshi [2015] for example), which reported, however, no significant information gain by expanding dataset more than core observable variables. The factor loading matrix of DFM, which is linear and only considers contemporaneous correlations, seems too restrictive to extract potentially valuable information from the big dataset.

2.4. TDVAE-DSGE with correction term

In the framework of TDVAE-DSGE described above, a law of motion (or transition model) of state variables is fully specified by the values of corresponding DSGE matrix. As discussed in the first section, this might be too restrictive to capture potentially non-stationary nature of the economic data. In other words, actual developments of economy (or true data generation process) could often deviate from the assumptions of a plain-vanilla linearized New Keynesian DSGE model. In order to address this problem, I choose to insert a neural-net-based correction term into the transition model of TDVAE-DSGE. Thus, the transition probability with a correction term is described as follows:

$$
\begin{aligned}
&\left[CT, \log \sigma_{z^T}\right] = D(z_{t-1}^S), \\
&\mu_{z^T} = F z_{t-1}^S + CT, \\
&z_t^T \sim p_T^{t|t-1} = N(\mu_{z^T}, \sigma_{z^T}).
\end{aligned}
$$

(5)

With its non-linear nature of a neural network, the correction term CT could help to capture flexibly the residual dynamics of state variables, which accounts for the movement of observed data not identified by the DSGE model. Hereafter, this modified structure is denoted as TDVAE-DSGE-CT.

3. Empirical Results

This section compares performances of the models introduced above and some benchmarks with regard to forecasting back-test. Precisely, following three versions of TDVAE-DSGEs and three benchmarks are tested on the equal-footing basis.

(1) TDVAE-DSGE-7V (7V): this version of TDVAE-DSGE uses only the seven core observation variables (i.e. production, consumption, investment, labor input, wage, price, interest rate) as inputs to construct belief states.

(2) TDVAE-DSGE-DR (DR): this is the data-rich version of TDVAE-DSGE where belief states compress a large number (more than 500) of economic and financial data in addition to the seven core variables. Decoder probability (i.e. observation likelihood) remains to be calculated with the seven core variables.

(3) TDVAE-DSGE-CT (CT): this version has a correction term in its transition model as explained in the section 2.4, where belief states are constructed from the big data as with the DR version.

(4) VAR: a simple reduced-form VAR (2) model, augmenting the core seven variables.

(5) PF-DSGE (PF): the estimated DSGE model by particle filter explained above. Forecasts are constructed by randomly sampled structural shocks for quarters ahead.

(6) LSTM: a standard LSTM network taking lagged values of the seven core variables to explain the current values of them, intended to capture a potentially non-linear low of motion.

In a back-test procedure, forecasts for real GDP by each model have been assessed, and all generated with no future information except for the use of revised data. Back-test periods are from 1Q2011 to 4Q2018, in which all models are re-estimated by every quarter, with the data that had been available at the release date of GDP statistics. The corresponding release dates to the values of each economic indicators are corrected from the sources, but for the periods of unavailable pasts, complemented by the largest days of delay in the recent known release dates.

For every quarters of back testing, forecasts up-to eight quarters ahead were generated after the re-estimation. All models take real GDP in the form of log-detrended level, so do their forecasts. In order to assess the performance of each model, Root Mean Squared Error (RMSE) is calculated for two, four and eight quarters-ahead forecasts. When evaluating forecasts of a level variable with some persistency, it is known that simple autoregressive model like AR(1) or even side-slide forecasts could record a high RMSE accuracy, which is so called echo-effects. The actual forecasts by PF-DSGE and VAR indeed exhibited the patterns of side-slide, in which their RMSEs are thought of too low with the existence of such echo-effects. Therefore, I choose to additionally test the performances in terms of quarter by quarter (qoq) changes. In this case, RMSEs up-to two, four or eight quarters ahead are calculated first at each period, and then averaged over all back-test periods. All those results are summarized in Table 1. Overall score of each model is evaluated by the ranking over four patterns of those RMSEs.

Table 1. Summary of the back-test results

RMSE 1Q2011 to 4Q2018

	PF	VAR	LSTM	TDVAE-7V	TDVAE-DR	TDVAE-CT
q+2	1.40	1.49	1.48	2.09	1.58	1.03
q+4	1.70	2.31	1.88	2.26	1.57	1.05
q+8	0.82	3.05	2.29	2.11	2.09	0.96
qoq_2q	0.75	0.71	0.63	0.55	0.49	0.48
qoq_4q	0.71	0.80	0.68	0.60	0.59	0.62
qoq_8q	0.68	0.74	0.65	0.63	0.64	0.64
Total Ranking	4	6	5	3	2	1

4. Discussions

4.1. Role of correction term

Assessing with the overall ranking, the best score was achieved by TDVAE-DSGE-CT (CT) as anticipated. This outperformance of CT against TDVAE-DSGE-DR (DR) suggests that the correction term constructed

through neural networks played an important role to capture a potentially non-linear dynamics (or underlying data generation process) of the state variables, corresponding to the structural shocks of macroeconomy under the DSGE restriction.

Comparing the actual filtered values of the seven state variables under CT and DR, some variables show acute differences in particular periods despite the considerable resemblances globally for most variables. Concretely, the productivity process under CT shows a peak and trough during the late 1980s and early 1990s (the Japan's financial bubble and burst), and also a much deeper correction in 2008-2009 (the Lehman shock), which are not appeared in the process under DR. The difference between CT and DR is also visible in the wage-markup process in most periods. In addition, the process of monetary policy shock has been broadly non-similar.

These may suggest the modeling of financial sector, labor market, and monetary policy in the plain-vanilla DSGE of Smets and Wouters [2007] are not adequate to explain the underlying data generation process of Japan's macroeconomy during the corresponding periods. Indeed, the importance of richer description of financial sector was also emphasized in Hasumi et.al [2018].

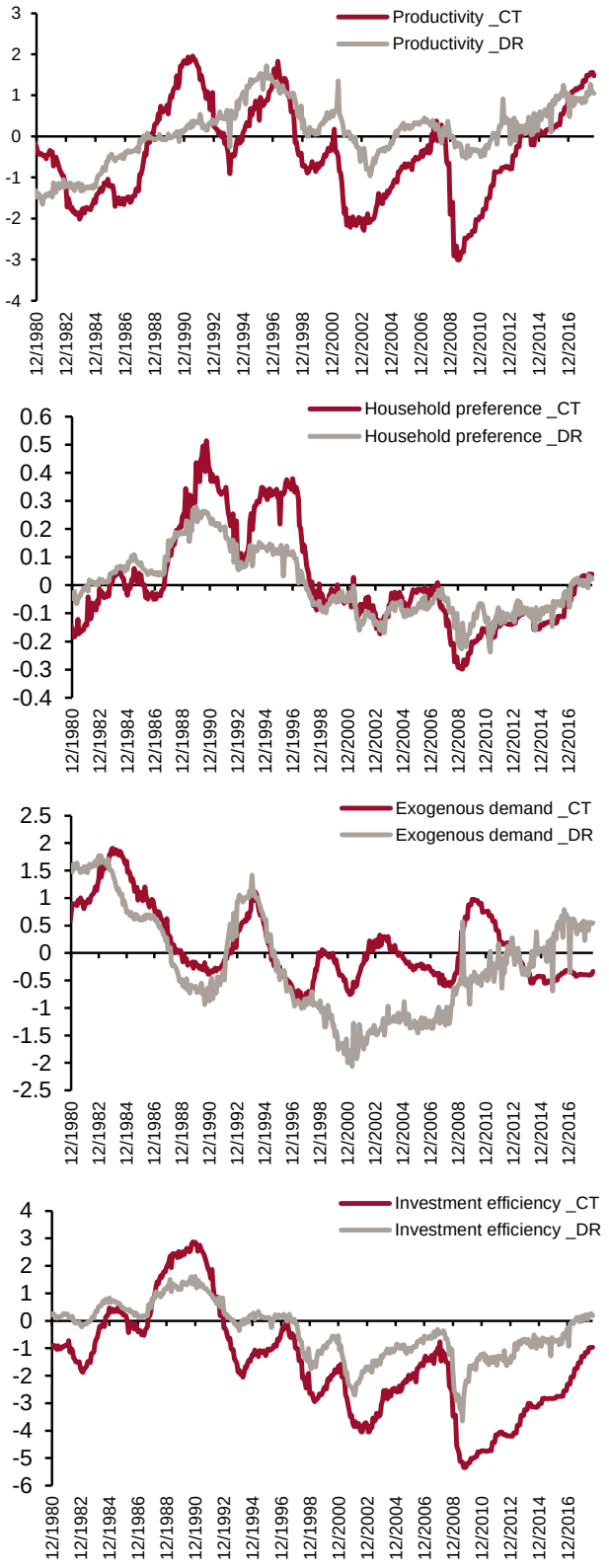
4.2. Data-rich beliefs

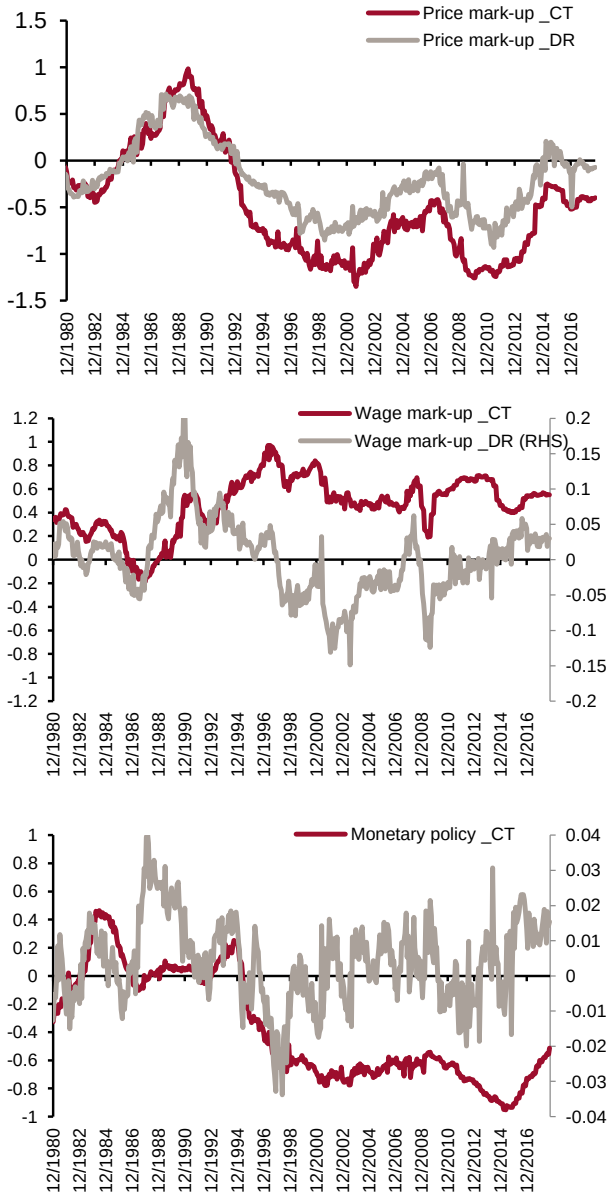
The better performance of TDVAE-DSGE-DR (DR) to TDVAE-DSGE-7V (7V) implies that the use of richer dataset to form belief states leads to more accurate inference of the state variables, and their dynamics. The 7V did not so significantly outperform the standard particle filter estimation of DSGE (PF) in the overall accuracy ranking, which suggest that the advantage of TDVAE methodology tends only materialized with a richer dataset. These results imply that non-linear flexible structure of LSTM could efficiently extract valuable information from large number of economic indicators, which is not contained in the seven core observable variables.

4.3. Theoretical restriction

Both the higher accuracy of TDVAE-DSGEs against LSTM and the same of PF-DSGE against VAR suggest that theoretical restrictions by a DSGE model certainly enhanced the out-of-sample forecasting performances of Japan's GDP, regardless of the type of models, namely, neural networks or linear autoregression.

Figure 1. Filtered values of the states by TDVAE-DSGE-CT vs. DR





5. Conclusion

This paper examined the possibility of combining a DSGE model and neural networks to supplement each other, with regard to out-of-sample forecasts for economic variables. The aim is to build a model with theoretical interpretability and state-of-the-art performance. The novel neural-net structure of TDVAE (Temporal Difference Variational Auto-Encoder) proposed by Gregor et.al [2019] enables to realize this idea. TDVAE virtually replicates a Gaussian stochastic state-space model by combination of neural networks. Because a DSGE model provides theoretical restrictions on the state transition and observation matrices of a linear state-space model, I choose to transplant those DSGE-oriented matrices into the formulations of state transition and observation probabilities in TDVAE. This TDVAE-DSGE approach certainly achieved the superior performance in the task of out-of-sample forecasts on Japan's real GDP during 1Q/2011 and 4Q/2018.

References

[1] Borovkova, S. and Tsiamas, I. (2019). An ensemble of LSTM neural networks for high‐frequency stock market classification. Journal of Forecasting, 2019;1-20. https://doi.org/10.1002/for.2585

[2] Buesing, L., Weber, T., Racaniere, S., Ali Eslami, S., Rezende, D., Reichert, D., Viola, F., Besse, F., Gregor, K., Hassabis, D., and Wierstra, D. (2018). Learning and querying fast generative models for reinforcement learning. arXiv preprint arXiv:1802.03006.

[3] Chen A., Leung M.T., and Daouk H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan Stock Index. Computers & Operations Research, 30: 901-923.

[4] Del Negro, M., Eusepi, S., Giannoni, M., Sbordone, A., Tambalotti, A., Cocci, M., Hasegawa, R., Linder, and M. (2013). The FRBNY DSGE model, Staff Reports 647. Federal Reserve Bank of New York.

[5] Rochelle, M., and Refet, G. (2010). How useful are estimated DSGE model forecasts for central bankers?. CEPR Discussion Papers, No 8158.

[6] Fischer, T., and Krauss, C. (2017).Deep learning with long short-term memory networks for financial market predictions. FAU Discussion Papers in Economics 11/2017.

[7] Gelfer, D.G. (2016). Financial crises and labor market dynamics: evidence from a data-rich DSGE model. Mimeo.

[8] Gregor, K., Papamakarios, G., Besse, F., Buesing, L., and Weber, T. (2019). Temporal Difference Variational Auto-Encoder. arXiv preprint arXiv:1806.03107.

[9] Hasumi et.al [2018]

[10] Hasumi, R., Iiboshi, H., Matsumae, T., and Nakamura, D. (2018). Does a financial accelerator improve forecasts during financial crises?: Evidence from Japan with Prediction Pool Methods. MPRA Paper, No. 85523.

[11] Herbst, E., and Schorfheide, F. (2012). Evaluating DSGE model forecast of comovements. Journal of Econometrics, 171(2), 152-166.

[12] Iiboshi, H., Matsumae, T., Namba, R., and Nishiyama, S. (2015). Estimating a DSGE model for Japan in a data-rich environment. Journal of the Japanese and International Economies, vol. 36(C), pages 25-55.

[13] Klein, P., (2000). Using the generalized schur form to solve a multivariate linear rational expectations model. Journal of Economic Dynamics and Control, 24, 1405–1423.

[14] Sims, C.A., (2001). Solving linear rational expectations models. Journal of Computational Economics, 20 (1–2), 1–20.

[15] Smets,F., and Wouters, R. (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. American Economic Review, 97 (3): 586-606. DOI: 10.1257/aer.97.3.586.

銘柄属性と市場情報を用いた深層学習株式リターン予測モデルの予測根拠の分析

Analyzing Basis of Stock Returns Prediction Model Based Deep Learning Using Stock Attributions and Market Information

小寺俊哉[1]　佐藤史仁[1]　坂地泰紀[2]　和泉潔[2]

Shunya Kodera[1], Fumihito Sato[1], Hiroki Sakaji[2], and Kiyoshi Izumi[2]

[1] 日興リサーチセンター株式会社　投資工学研究所
[1]Institute of Investment Technology, Nikko Research Center, Inc.
[2] 東京大学大学院工学系研究科　システム創成学専攻
[2]Department of Systems Innovation, School of Engineering, The University of Tokyo

Abstract: 近年，個別株式のリターン予測において，様々なファクターの中から予測に有用な特徴量を自動で抽出することのできる深層学習技術の応用研究がなされている．しかしながら，深層学習は計算の過程が複雑で，人間にはその予測根拠の把握が難しく，意思決定に理由が求められる実務での利用において，解釈の困難さが課題視されることがある．一方，深層学習の解釈手法についても研究が行われており，深層学習において研究が盛んな画像分類等のタスクだけでなく株価指数や株式個別銘柄等の資産価格リターン予測を行う深層学習モデルに対しても解釈を行う研究が行われている．本稿では，モデルの解釈に焦点を当て，個別銘柄のリターン予測をタスクとした深層学習モデルについて，LRP と呼ばれる深層学習の解釈手法を用いて，各入力値の寄与度を個別銘柄レベルで確認した．さらに，深層学習モデルの入力値に個別銘柄属性だけでなくマーケット指数等の市場情報を用いることで，銘柄属性と市場トレンドの 2 つの側面での分析を行った．

1　はじめに

近年，ビッグデータと呼ばれる大規模で多種多様なデータの活用や計算機の大規模並列処理能力の向上を背景に，与えられたデータからパターンを抽出し，分類や予測を行う機械学習の技術が目覚ましい発展を遂げている．特に，分類や予測等，対象とするタスクに有効な特徴量を人手で与えることなく，自ら選択, 抽出することのできる深層学習の技術が，様々な分野で注目を浴びている．

資産運用業界等でも，ポートフォリオマネジメントにおけるリスク管理や投資戦略等に対し，深層学習を用いた研究がなされている．例えば，株価指数や株式個別銘柄のリターン等の予測[11][23]，それを応用したポートフォリオ選択[3][13]，リスクヘッジに関する予測[7], ボラティリティ予測[14][20]等が挙げられる．その中でもリターン予測は，投資戦略だけでなくポートフォリオ構築やリスク管理にも活用できる非常に関心の高いテーマとなっている．

ところで，ファイナンス分野の研究において，株式等の資産のリターンを説明するファクターモデルがいくつも提案されている．最も基本的なモデルである資本資産評価モデル（CAPM）では，個別資産のリターンは，市場全体のマーケットポートフォリオに対する感応度 β で特徴づけられている[10][15]. そして，株式の個別銘柄のリターンに対しては，β 以外に，企業規模や簿価時価比率を説明力のある変数として組み入れた Fama-French の 3 ファクターモデル[4]や，さらにモメンタムを組み入れた Carhart の 4 ファクターモデル[2]，その他，財務変数等を組み入れた様々なマルチファクターモデルが提案されている[5][6][8]. このような個別銘柄のリターンに対して一定の説明力があると報告された変数は 300 以上あるものの，どの変数がリターンを説明する本質的な要因であるか，もしくはそれらをどのように組み合わせてリターンを説明するモデルとするか，未だ明確な答えはない．また，運用実務において，個別銘柄間の相対的な投資魅力度を定量的に推計するために，いくつかの銘柄属性等の変数をファクターとし

て，クロスセクションで各銘柄のリターンを説明するマルチファクターモデルが利用されることがある．しかし，多数のファクターの膨大な組み合わせから有効なものをどう選択するか，また，ファクター間やファクターとリターンとの非線形の関係をどのように取り込むかが課題となる場合が多い．このような課題に対して，深層学習は，与えられたデータから有効な特徴量を自動で抽出し，線形・非線形なモデルの構築が行える手法として高い親和性が期待される．

しかしながら，深層学習が特徴量の自動抽出や非線形の複雑な表現を扱える一方で，入力と出力間にどのような関係性を学習したか人間が理解できる形で確認することは難しく，ブラックボックス化してしまう問題がある．深層学習を銘柄選定等の投資運用実務に用いる場合，この問題によって投資判断の根拠を示しにくく，投資信託等の運用マネージャーの説明責任や，投資家の投資への納得性に対して障壁になることが少なくない．これらは，深層学習技術の実務への応用に対する重要な課題の一つだろう．

一方で，深層学習の解釈手法[1]に関する研究も盛んに行われており，様々な手法が提案されている[9][18]．例えば LRP(Layer-Wise Relevance Propagation)[2][1]と呼ばれる手法は，深層学習のモデル構造に手を加えず，モデルの出力値からその計算を遡ることで各入力値が出力に対してどの程度貢献したかを計算することができる．この技術を用いて，深層学習での画像分類やテキストの極性判定において，どの入力部分が貢献したかを分析した研究が報告されている[19][21]．

金融分野においても，深層学習モデルの予測や推論の根拠の理解に関する研究は注目度が高く，深層学習を用いてリターン予測や株式選択を行う際に，各入力変数の貢献度を分析して予測結果の解釈性を高める研究が盛んに行われている。Shi ら[16]は深層学習を用いて株式を選択する手法を提案し，多変量解析法の一つである MDS（多次元尺度構成法）を用いて推論根拠の可視化を行った．Wu ら[17]は GRU（Gated Recurrent Unit）ネットワークと呼ばれる深層学習の手法を用いて，過去の価格データから将来の株価リターンを予測し，予測結果をユーザに解釈しやすい決定木の形に変換する GRU-Tree という解釈可能なモデルを提案した．塩野[22]は，多数の経済金融時系列を入力とし，資産価格リターンの予想を行う CNN（畳み込みニューラルネットワーク）を，マクロ経済の理論モデルとのマルチタスク学習により過学習を緩和する方法を提案した．その際に，LRP を応用することで，マクロ経済モデルでの経済ファクターと資産価格リターン予測における共通因子を特定し，モデルの解釈性を高めた．Nakagawa ら[12]は，深層学習による株式個別のリターン予測モデルに LRP を適用し各入力値の影響を分析する手法を提案した．これらの研究に対し，本稿では，株式個別銘柄のリターン予測をタスクとした深層学習モデルの出力の解釈に焦点を当てた分析を行った．深層学習の解釈性を高める手法は[12][22]と同様に LRP を用いた．予測根拠の解釈対象の深層学習モデルは，入力データに ROE 等の個別銘柄の属性であるファクター値の他，マーケット指数等の市場情報を用い，ファンダメンタルズと市場全体のトレンドの両方を考慮したモデルとした．このことにより，関連研究にはない銘柄属性と市場トレンドの 2 つの側面での貢献度分析が可能となる．具体的には，深層学習が，入力値とした代表的な各銘柄属性に対して，どのような特性を持つ銘柄のリターンを相対的に高く（又は低く）予測したかを，予測リターン順を基にしたロングポートフォリオとショートポートフォリオの平均 LRP 貢献度や，入力ファクターとその LRP 貢献度の散布図から確認した他，個別銘柄レベルでの各属性の貢献度を確認した．また，マーケット指数等の LRP 貢献度の時系列の変化や，特定の局面において特に市場情報に影響を受けていた銘柄の属性について分析を行った．

2　実験設定

2.1　リターン予測の深層学習モデル設定

まず，LRP による予測根拠の解釈を行う学習済みモデルとして，本稿では，クロスセクションでの株式個別銘柄の相対的な投資魅力度を，時価総額や自己資本比率等の銘柄属性等から予測する深層学習モデルを構築した．深層学習モデルは，多層パーセプトロンを用いた．また，各月末ごとにモデルを構築し 1 ヶ月後を予測するローリング方式を想定した．教師データは，その月末から翌月末までのリターンとし，入力データは，各月末時点で入手可能なデータから作成した個別銘柄の属性である各ファクター

[1] 説明可能な AI 技術及びその研究分野の事を XAI（Explainable AI）と呼び，内閣府の人工知能技術戦略会議の人工知能技術戦略実行計画（https://www8.cao.go.jp/cstp/tyousakai/jinkochino/keikaku.pdf）や総務省の AI ネットワーク社会推進会議の 2018 年報告書（http://www.soumu.go.jp/main_content/000564147.pdf）にも明記されている人工知能研究において重要なテーマの 1 つである．

[2] LRP は，深層学習モデルにおいて通常通りにデータを入力し順伝播させた後に，レイヤー間の関係性を逆伝播することで入力と出力の関係性を求める手法で，サンプルごとに出力値に対する各入力値の貢献度を計算することができる．

値や東証 33 業種のダミー変数の他, 月のダミー変数やマーケット指数等，計 126 変数（各変数は表 3，4 参照）を用いた．なお，各データは必要に応じて標準化や正規化の処理を行った．学習データには，各月末時点で東証一部に上場している銘柄をユニバースとし，計算可能な入力データと教師データ過去 36 ヶ月分を用いた．学習したモデルによって予測される各銘柄の 1 ヶ月リターンの順位付けを行い，これを相対的投資魅力度とした．なお，モデルのハイパーパラメータは予測日によらず共通の値（表 5 参照）を用いた．

2.2 パフォーマンス検証

各月末に構築した深層学習モデルで予想される投資魅力度の評価として，投資魅力度をもとにロングとショートのポジションを構築し，月次リバランスでロングショートポートフォリオを運用した場合のパフォーマンスを計測した．ユニバースは東証一部上場銘柄とし, 投資魅力度の高い順に 10 分位ポートフォリオを等ウェイトで構築，最高分位をロング，最低分位をショートのポートフォリオとした．各ポートフォリオの平均リターンの差であるスプレッドリターンをロングショートポートフォリオのリターンとして 2005 年 12 月末から 2019 年 12 月末まで計測し，各種パフォーマンス値を表 1 に示した．

ロングショートポートフォリオのリターン/リスク比は 1.63 であった．また，各パフォーマンス値について，ロングポートフォリオは，TOPIX 対比でアウトパフォームしたのに対し，ショートポートフォリオはアンダーパフォームとなった．実際の運用では，売買コストやキャパシティ，投資可能な営業日数等の制約があることに注意が必要ではあるが，これらの結果から，深層学習モデルによるリターン予測を投資戦略等へ応用することで，一定のパフォーマンスを得られることが期待できるだろう．

表 1. ポートフォリオの各種パフォーマンス値

	ロング	ショート	ロング ショート	TOPIX
リターン	15.91%	-5.20%	20.72%	2.31%
リスク	19.67%	21.52%	12.75%	17.72%
リターン/ リスク比	0.81	-0.24	1.63	0.13
最大ドロー ダウン	-44.18%	-74.42%	-22.38%	-56.23%

（注）リターン及びリスクは年率の値を示す．参考として TOPIX（配当込み）も示す．

3 予測根拠の分析

LRP 貢献度は，深層学習モデルによる個別銘柄ごとの予測リターンの計算に用いられた全ての入力値について算出される．各 LRP 貢献度はプラスに大きいほど，その入力のファクター値が出力の予測リターンに対してプラスに大きく寄与し，マイナスに大きいほど，マイナスに大きく寄与することを表す．

本稿では，入力データの内，代表的な銘柄属性ファクターの LRP 貢献度に対し，予測リターンが高い銘柄と低い銘柄間での傾向等を確認することで，予測リターンの根拠の分析を行った．また，2019 年 6 月末の予測日における深層学習の予測結果を対象とし，予測リターンが高い銘柄を前述のロングポートフォリオの銘柄，低い銘柄をショートポートフォリオの銘柄として分析を行った．さらに，市場情報について，ロングとショートのポートフォリオにおける平均 LRP 貢献度を算出し，その時系列変化や，特定の局面において，予測リターンに対して市場情報がポジティブ（又はネガティブ）に影響した銘柄群の特性について調べた．

3.1 個別銘柄の予測に対する根拠の提示

まず，個別銘柄ごとにどのファクターが予測に寄与しているのかを確認した．表 2 は，ロングとショートのポートフォリオからそれぞれ 5 銘柄ずつ選択した計 10 銘柄の予測リターンとファクター値及び LRP 貢献度を示している. 予想経常利益修正～予想売上高成長率の表中数値は深層学習モデルに入力された各ファクター値，その色は LRP 貢献度を示す．LRP 貢献度が 0 の場合に白色, プラスに大きいほど赤色が濃く，マイナスに大きいほど青色が濃くなるように色を付与した．

銘柄ごとの予測リターンに対するファクターの LRP 貢献度を見ると，例えば<4502>の場合では，予想経常利益修正や予想益利回り, 時価総額, 予想 ROE, 予想売上高成長率の色が濃く，リターン予測に大きく寄与していたことが分かる．また，これらファクターについて，予想売上高成長率は赤色であるためプラス方向，それ以外は青色であるためマイナス方向の寄与であることが分かる．ファクター値と合わせて見ると, 予測リターンに対して, 売上の拡大（予想売上高成長率のファクター値がプラス）に関してプラス方向に評価（LRP 貢献度が赤色）されているものの，収益性の低下（予想 ROE，予想経常利益修正のファクター値がマイナス）や割高（予想益利回

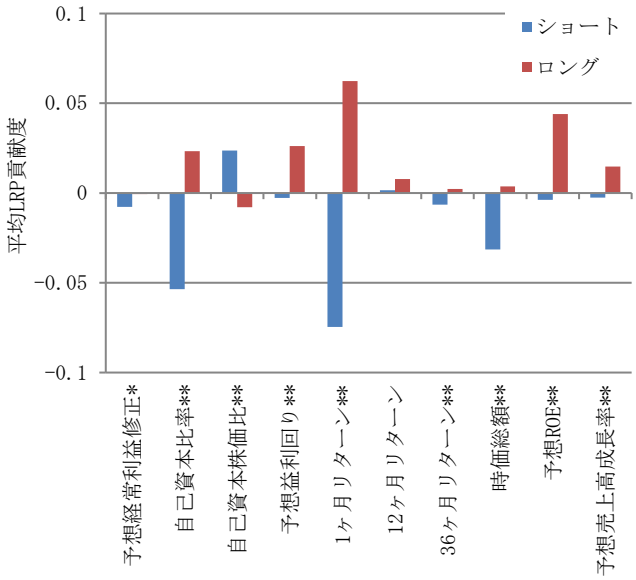
表 2. 個別銘柄の予測リターンと各ファクター値及び LRP 貢献度

ポートフォリオ	銘柄	予測リターン	数値：入力変数値 / 色：LRP 貢献度（高[赤]-0[白]-低[青]）									
			予想経常利益修正	自己資本比率	自己資本株価比	予想益利回り	1ヶ月リターン	12ヶ月リターン	36ヶ月リターン	時価総額	予想 ROE	予想売上高成長率
ショート	4502	-0.40	-3.14	-0.66	-0.33	-3.02	0.27	-0.03	-0.77	2.90	-2.52	3.74
	6326	-0.28	-0.02	-0.23	-0.65	-0.20	0.97	0.65	0.15	2.28	0.23	0.25
	2503	-0.27	-1.07	-0.54	-0.88	-0.99	-0.36	-0.08	0.22	2.25	-0.37	-0.09
	9202	-0.25	-0.10	-0.49	-0.30	0.26	-0.65	0.41	-0.04	1.92	0.08	0.01
	7974	-0.25	-0.06	1.54	-1.09	-0.82	-0.05	1.17	1.94	2.81	0.66	0.26
ロング	3769	0.15	-0.02	-1.42	-1.38	-1.43	-0.13	1.34	1.69	1.42	1.80	2.40
	4188	0.15	-2.32	-1.26	0.13	1.60	0.51	-0.10	0.76	1.86	0.44	-0.04
	4684	0.17	0.09	1.82	-1.20	-1.05	-1.49	2.27	1.37	1.91	0.83	0.41
	3765	0.18	1.01	1.22	-1.10	-0.29	-2.20	1.55	-0.38	1.00	2.18	-0.14
	6770	0.27	-0.44	0.14	-0.25	0.09	-0.22	-0.96	-0.68	1.21	-0.10	-0.27

りのファクター値がマイナス）によるマイナス評価（LRP 貢献度が青色)の寄与がより大きかったため，投資魅力度が低く判定されたと推察することができる．同様に解釈すると，<3765>は，ファクター値と LRP 貢献度を合わせて見ると，予測リターンに対して，割高気味（自己資本株価比のファクター値がマイナス，12 ヶ月リターンがプラス）であることがマイナスに評価（LRP 貢献度が青色）されるものの，収益性（予想 ROE のファクター値がプラス）や安全性（自己資本比率のファクター値がプラス）についてはプラス評価（LRP 貢献度が赤色）を受け，さらに直近の株価下落（1 ヶ月リターンのファクター値がマイナス）に対するリバウンドが見込めるため投資魅力度が高く判定されたと推察することができる．

3.2 ファクター効果の分析

次に，投資魅力度の高い（又は低い）銘柄群では，どのファクターがその予測に寄与している傾向があるのかを確認するため，個別銘柄ごとに算出された LRP 貢献度について，ロングとショートのポートフォリオでの各ファクターの平均値を算出した．また，ファクターごとのポートフォリオ間での LRP 貢献度の平均に有意な差があるかの検定も行った．その結果を図 1 に示す．ロングポートフォリオでは，1 ヶ月リターン，予想 ROE，予想益利回り，自己資本比率がプラス予測に寄与していた．ショートポートフォリオでは，1 ヶ月リターン，自己資本比率，時価総額がマイナス予測に寄与していた．これらは全て 1%



（注）各ファクターにおける平均の差の検定において，*は 5%，**は 1%有意水準を示す．

図 1. ポートフォリオの平均 LRP 貢献度

有意水準となった．また，1 ヶ月リターンのように，ロング側でもショート側でも予測に対する寄与が大きいファクターがある一方で，予想 ROE や時価総額のように，ロング側かショート側かどちらか一方の予測に対して寄与するファクターもあった．入力データの分布にもよるが，投資魅力度の高い（低い）予測に対して，深層学習モデルがどのファクターに注目しているかの傾向を知る手掛かりになるだろう．

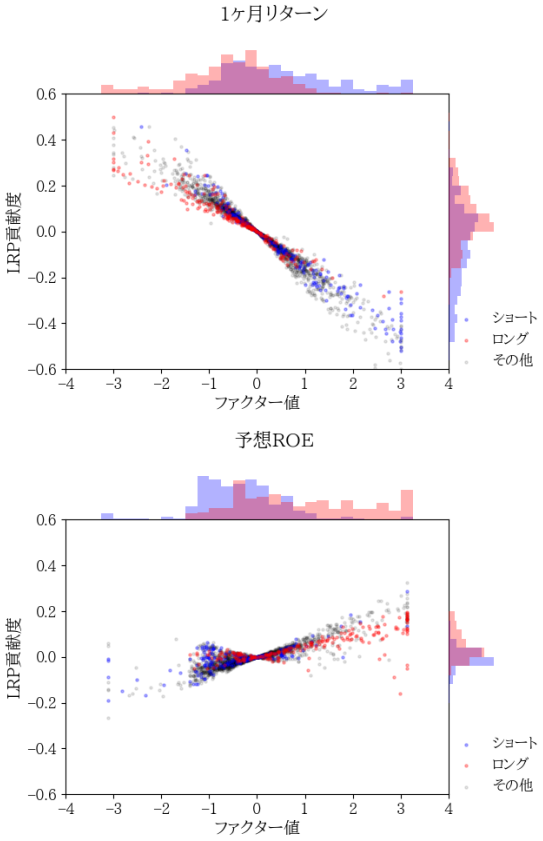


図 2. ファクター値と LRP 貢献度の関係

図 1 で, 特に 1 ヶ月リターンや予想 ROE 等のファクターにおいてポートフォリオ間での平均 LRP 貢献度の差が大きいことが分かった．しかし，この情報だけでは，どのファクターが予測リターンに寄与している傾向にあるかは分かるが，ファクターの値の大小と予測リターンの関係性は確認することができない．つまり，ファクターの値と LRP 貢献度の関係性や，これらと予測リターンの関係性で表わされる個別銘柄レベルでの全体的な傾向を確認することができない．そこで，銘柄のファクター値と LRP 貢献度の散布図と各変数のロングとショートのポートフォリオ別のヒストグラムを用いて確認した．例として, 図 2 に 1 ヶ月リターンと予想 ROE での結果を示す．1 ヶ月リターンではファクター値と LRP 貢献度は全体として負の相関関係にあり，特にロングポートフォリオの銘柄ではファクター値はマイナス，LRP 貢献度はプラスに裾野が広がっていた．ショートポートフォリオの銘柄では逆にファクター値はプラス，LRP 貢献度はマイナスに裾野が広がっている傾向が見られた．一方，予想 ROE ではファクター値がプラスの範囲ではファクター値と LRP 貢献度は正の相関があるが，ファクター値がマイナスの範囲では明確な相関は見られず，ロングポートフォリオの銘柄では LRP 貢献度の分布はプラス側に寄っているが，ショートポートフォリオの銘柄では LRP 貢献度の分布は 0 付近に集中していた．1 ヶ月リターンにおいて，ファクター値と LRP 貢献度には負の相関があることから，この深層学習モデルは，この月の東証一部市場の全体の傾向として，短期リバーサル効果が働くことを予想していると解釈することができるだろう．また，予想 ROE において，高収益性で見たクオリティー効果が働くことを予想していると解釈できるだろう．さらに，ロングやショートのポートフォリオの銘柄に対し，LRP 貢献度が 0 付近のものについて，そのファクター以外のファクターが予測リターンに寄与していると解釈できるだろう．

興味深い点として，銘柄間でファクター値と LRP 貢献度に相関がある中でも，同じファクター値の水準でありながら LRP 貢献度に正負が異なるようなばらつきがある点や，ファクター値が正の範囲と負の範囲でファクター値と LRP 貢献度の関係性が異なる傾向が見えたことが挙げられる．これらは，深層学習の予測に他のファクターとの関係性や非線形性が考慮されていた結果であるかもしれない．

3.3 市場情報を用いた分析

本稿で分析対象とした深層学習モデルでは，銘柄属性だけでなく，TOPIX や為替といった市場情報も入力値に用いた．これらの情報の予測リターンへの寄与を確認することで，深層学習モデルが市場情報をどう取り入れて予測したかについて分析した．

図 3 では，4 つの市場情報についてロングとショートの各ポートフォリオでの LRP 貢献度の平均値の差を時系列で確認した．市場情報は，クロスセクションでは銘柄間で同じ値の入力であるものの，時期によってロングとショートのポートフォリオ間で異なる傾向が見られた．例えば，TOPIX 変化率は 2008 年 10 月や 2011 年 3 月，2012 年 2 月，2016 年 7 月，2019 年 10 月にロングとショートのポートフォリオ間で LRP 貢献度の平均の差が大きかった．その他にも，日本国債金利 10 年は 2011 年 3 月，ドル/円は，2013 年 12 月や，2014 年 8 月，2019 年 3 月で差が大きかった．一方で，他の指標と比べ VIX では目立った動きは見られなかった．

次に, TOPIX 変化率において LRP 貢献度のポートフォリオ間の差が大きかった 2008 年 10 月に注目した．2008 年 10 月は，2008 年 9 月に起きたリーマンショックによる金融危機の影響で TOPIX は大きく下落した．深層学習によるリターン予測において，この TOPIX の下落をポジティブやネガティブに捉えられた銘柄群にそれぞれどのような特性があるか

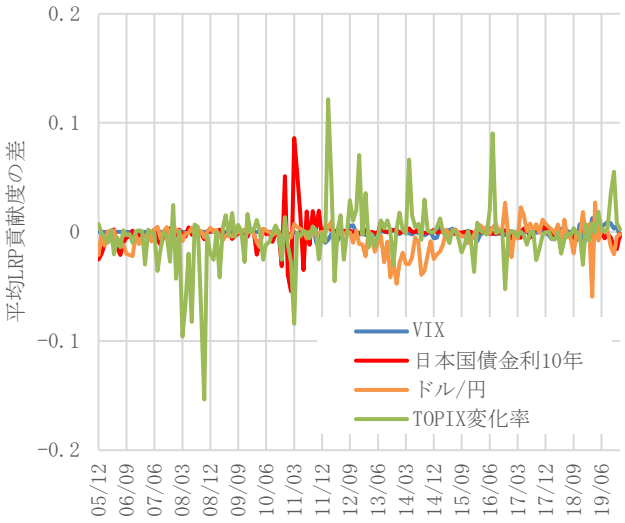


図 3. ロングとショートポートフォリオ間での各市場情報の平均 LRP 貢献度の差

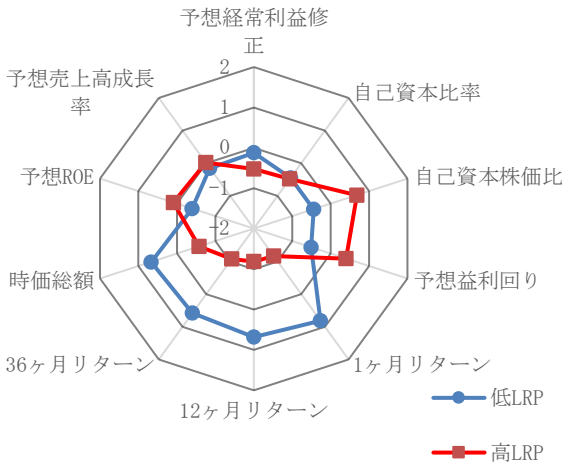


図 4. 2008 年 10 月末における TOPIX 変化率の LRP による分位の各平均ファクター値

を確認した．図 4 では，2008 年 10 月末を予測日とした際の TOPIX 変化率の LRP 貢献度で 10 分位し，最も高分位と低分位の銘柄をそれぞれ高 LRP，低 LRP ポートフォリオとし，代表的な 10 個の銘柄属性についてファクター値の平均を示した．その結果，TOPIX の下落に対し深層学習の予測がポジティブな影響を受けている銘柄の属性は，相対的に自己資本株価比や予想益利回りが高く，1 ヶ月リターン，12 ヶ月リターン，36 ヶ月リターン，時価総額が低かった．反対にネガティブな影響を受けている銘柄の属性は，相対的に自己資本株価比や予想益利回りが低く，1 ヶ月リターン，12 ヶ月リターン，36 ヶ月リターン，時価総額が高いという結果であった．この結果から，深層学習モデルは TOPIX の下落が予測リターンに対してポジティブかネガティブか銘柄属性によって異なるよう学習していると解釈できるだろう．

4　まとめ

本稿では，LRP を用いて深層学習による個別銘柄の予測リターンに対する根拠の分析を行った．

表 2 では，銘柄ごとにどのようなファクターが予測に寄与しているかの確認を行った．例えば，<4502>は収益性の低下や割高であることが要因として投資魅力度が低く判定され，<3765>は収益性や安全性が高いことや，直近株価の下落に対するリバウンドを見込んで投資魅力度が高く判定されたという解釈を得ることができた．

また，図 1 や図 2 から，投資魅力度の高い（又は低い）銘柄群で予測リターンに対し，どのようなファクターの寄与が大きい傾向にあるか，市場全体での予測リターンに対するファクター効果として，個々の銘柄の予測におけるファクター値と LRP 貢献度の関係について分析した．具体的には，対象とした 2019 年 6 月末を予測日として構築した深層学習モデルが，東証一部の銘柄に対し，1 ヶ月リターンについて短期のリバーサル効果を予測していることや，予想 ROE では，高収益性で見たクオリティー効果を予測しているとの解釈結果を得ることができた．

図 3 では，入力値に用いた市場情報の内 4 つについて，時系列でロングとショートのポートフォリオ間の平均 LRP 貢献度の差を分析した．市場情報は，個別銘柄間では，入力値としては同じ値だが，出力の予測リターンに対して異なった寄与を持つことが確認できた．また，2008 年 10 月の金融危機による下落局面において，TOPIX の下落情報をポジティブ（又はネガティブ）に捉える銘柄の属性について図 4 で確認を行った．

5　おわりに

本稿では，LRP を用いて個別株式のリターン予測をタスクとした深層学習モデルに対して，ファクターごとの予測リターンに対する貢献度を算出し，その予測根拠の分析を行った．その結果，LRP 貢献度を測る事で深層学習モデルが各銘柄の予測において，どのファクターに着目してリターンの予測を行っているか，ファクターごとに全体的な傾向として予測リターンにどのように影響しているかを確認することができた．本稿で用いた手法により，ファクターとリターンの線形的な関係だけでなく，深層学習の特徴である他の変数との関係や非線形性を考慮したファクターの効果を確認することができた．

また，実際に予測根拠を解釈する場合，図 1 で概観を確認し，図 2 で注目すべきファクターの貢献度との関係性を調べ，最終的には表 2 の個別銘柄での貢献度を確認する等の手順で，ブラックボックスであった予測に対する根拠の手がかりが示せるだろう．

さらに，図 3，図 4 のように銘柄の固有属性であるファクターの他，クロスセクションで銘柄間共通の値となる市場情報についても寄与度分析を行うことで，市場環境に応じてどのような情報が予測に用いられるか確認することが可能となる．このように，銘柄ごとの市場環境の影響を分析することで，モデルのより詳細な解釈に繋がるだろう．

参考文献

[1] Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., and Samek, W.: On Pixel-Wise Explanations for Non-Linear Classifier Decisions by Layer-Wise Relevance Propagation, *PLoS ONE*, 10(7): e0130140, (2015)

[2] Carhart, M. M.: On Persistence in Mutual Fund Performance, *The Journal of Finance*, Vol. 52, No. 1, pp. 57-82, (1997)

[3] Chakravorty, G., Awasthi, A., and Silva, B. D.: Deep Learning for Global Tactical Asset Allocation, *Available at SSRN* 3242432, (2018)

[4] Fama, E. F., and French, K. R.: Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics*, Vol. 33, No. 1, pp. 3-56, (1993)

[5] Fama, E. F., and French, K. R.: A five-factor asset pricing model, *Journal of Financial Economics*, Vol. 116, No. 1, pp. 1-22, (2015)

[6] Fama, E. F., and French, K. R.: Choosing factors, *Journal of Financial Economics*, Vol. 128, No. 2, pp. 234-252, (2018)

[7] Fécamp, S., Mikael, J., and Warin, X.: Risk management with machine-learning-based algorithms, *arXiv*: 1902.05287v3, (2019)

[8] Hou, K., Xue, C., and Zhang, L.: Digesting Anomalies: An Investment Approach, *The Review of Financial Studies*, Vol. 28, No. 3, pp. 650-705, (2015)

[9] Ito, T., Tsubouchi, K., Sakaji, H., Izumi, K., and Yamashita, T.: CSNN: Contextual Sentiment Neural Network, *19th IEEE International Conference on Data Mining (ICDM)*, pp.1126-1131, (2019)

[10] Lintner, J.: The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets, *The Review of Economics and Statistics*, Vol. 47, No. 1, pp. 13-37, (1965)

[11] Moghaddam, A. H., Moghaddam, M. H., and Esfandyari M.: Stock market index prediction using artificial neural network, *Journal of Economics, Finance and Administrative Science*, Vol. 21, No. 41, pp. 89-93, (2016)

[12] Nakagawa, K., Uchida, T., and Aoshima, T.: Deep Factor Model, *ECML PKDD 2018 Workshops*, pp. 37-50, (2018)

[13] Obeidat, S., Shapiro, D., Lemay, M., MacPherson, M. K., and Bolic, M.: Adaptive Portfolio Asset Allocation Optimization with Deep Learning, *International Journal On Advances in Intelligent Systems*, Vol. 11, No. 1 & 2, (2018)

[14] Sardelich, M., and Manandhar, S.: Multimodal deep learning for short-term stock volatility prediction, *arXiv*: 1812.10479v1, (2018)

[15] Sharpe, W. F.: Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *The Journal of Finance*, Vol. 19, No. 3, pp. 425-442, (1964)

[16] Shi, L., Teng, Z., Wang, L., Zhang, Y., and Binder, A.: DeepClue: Visual Interpretation of Text-based Deep Stock Prediction, *IEEE Transactions on Knowledge and Data Engineering*, Vol.31, No.6, pp.1094-1108, (2018)

[17] Wu, W., Wang, Y., Fu, J., Yan, J., Wang, Z., Liu, J., Wang, W., and Wang, X.: Preliminary Study on Interpreting Stock Price Forecasting Based on Tree Regularization of GRU, *International Conference of Pioneering Computer Scientists, Engineers and Educators*, pp.476-487, (2019)

[18] Zeiler, M. D., and Fergus, R.: Visualizing and Understanding Convolutional Networks, *European Conference on Computer Vision - ECCV 2014*, pp. 818-833, (2014)

[19] 伊藤友貴, 坂地泰紀, 和泉潔: 深層学習を用いた経済テキスト可視化の検証, *第 20 回人工知能学会金融情報学研究会*, pp. 61-66, (2018)

[20] 五島圭一, 高橋大志, 山田哲也: 自然言語処理による景況感ニュース指数の構築とボラティリティ予測への応用, *日本銀行金融研究所ディスカッション・ペーパー・シリーズ*, 2019-J-3, (2019)

[21] 酒井博貴, 亀谷由隆, 曽田尚宏, 有江浩明: LRP 法による畳み込みニューラルネットワークにおける中間層の振る舞いの可視化, *人工知能学会全国大会論文集 JSAI 2018*, 2A2-04, (2018)

[22] 塩野剛志: アセット・リターン予測 AI とマクロ経済理論の融合―マルチタスク学習による正則化と識別―, *第 21 回人工知能学会金融情報学研究会*, pp.33-40, (2018)

[23] 宮崎邦洋, 松尾豊: Deep Learning を用いた株価予測の分析, *人工知能学会全国大会論文集 JSAI2017*, 2D3-OS-19a-3, (2017)

Appendix

表 3. 個別銘柄の属性値

入力変数名				
実績売上高 成長率	実績配当利回り	予想 EBITDA/EV 比率	6 ヶ月リターン	実績益利回り
予想売上高 成長率	予想配当利回り	時価総額	12 ヶ月リターン	予想益利回り
実績経常利益 成長率	自己資本株価比	自己資本	36 ヶ月リターン	売上高営業 利益率
予想経常利益 成長率	予想益利回り 変化幅	売上高	β 値	自己資本比率
実績純利益 成長率	時系列基準化 益利回り	出来高回転率	倒産確率	負債比率
予想純利益 成長率	予想キャッシュフロー 株価比	1 ヶ月リターン	予想経常利益 修正	純債務比率
予想 ROE	予想売上高 株価比	3 ヶ月リターン		

表 4. マーケット指数等

	生値	1 ヶ月 変化率	1 ヶ月 差
TOPIX	○	○	-
日経平均	○	○	-
NY ダウ	○	○	-
S&P500	○	○	-
DAX	○	○	-
FTSE	○	○	-
ドル/円	○	○	-
ユーロ/円	○	○	-
日本国債金利 2 年	○	-	○
日本国債金利 5 年	○	-	○
日本国債金利 10 年	○	-	○
日本国債金利 20 年	○	-	○

	生値	1 ヶ月 変化率	1 ヶ月 差
日本国債金利 30 年	○	-	○
米国債金利 2 年	○	-	○
米国債金利 5 年	○	-	○
米国債金利 10 年	○	-	○
米国債金利 30 年	○	-	○
独国債金利 10 年	○	-	○
英国債金利 10 年	○	-	○
日経平均 VI	○	○	-
VIX	○	○	-
CRB 指数	○	○	-
WTI 原油先物	○	○	-
金先物	○	○	-

（注）表中○のものを入力変数として用いる.

表 5. 深層学習のハイパーパラメータ

ハイパーパラメータ		値
入力層	ユニット数	126
隠れ層 1	ユニット数	64
	活性化関数	tanh
	ドロップアウト	0.4
隠れ層 2	ユニット数	64
	活性化関数	tanh
	ドロップアウト	0.4
隠れ層 3	ユニット数	32
	活性化関数	tanh
	ドロップアウト	0.4
出力層	ユニット数	1
	活性化関数	linear
損失関数		二乗誤差
最適化アルゴリズム		Adam
ミニバッチサイズ / エポック数		300 / 20

変分自己符号化器を用いたネットワーク埋め込みと金融ネットワークへの応用

Network Embedding using Variational Autoencoder and its Application to Financial Networks

川上 雄大 [1*]　　江口 浩二 [2†]
Yuta Kawakami[1]　Koji Eguchi[2]

[1] 神戸大学 大学院 システム情報学研究科
[1] Graduate School of System Informatics, Kobe University
[2] 広島大学 情報科学部
[2] School of Informatics and Data Science, Hiroshima University

Abstract: In this paper, we present low-dimensional embedding methods for interbank transaction networks. To address one important problem: how to obtain latent representations that well capture the structual properties of a given directed network, we propose a new network embedding model, Co-Variational Autoencoder (Co-VAE). Co-VAE simultaneously learns network embedding focusing on the links going into each node and that focusing on the links coming out of each node, attempting to reproduce the original adjacency matrix. Thereby, we can learn the Co-VAE network embedding model, simultaneously capturing both the latent representations of lender patterns and those borrower patterns. Using both latent representations, we can predict interest rates of interbank transactions.

1　はじめに

現実世界には, 様々な関係データが数多く存在しており, これらはノード (頂点) とリンク (辺) で構成されるネットワーク構造で表現することができる. このようなネットワークから, 構造的特徴をよく表した潜在表現を獲得することがネットワーク分析の課題の一つとしてあげられる.

本稿では, 自己符号化器 (Autoencoder: AE) [1] の一種である変分自己符号化器 (Variational Autoencoder: VAE) [2] に着目したネットワーク分析を試みる. 自己符号化器は, 出力が入力と等しくなるように, 符号化器と復号化器の学習を行うモデルであり, 入力データの次元数を圧縮した潜在表現を獲得することができる. この潜在表現は, データの余分な情報やノイズを取り除き, 重要な情報を残した高密度な表現であるといえる. 本稿では, データの特徴をより良く表現する潜在表現を獲得することを目的としており, これが実現できれば, ノード間のリンクの有無の予測やその属性値の予測などのタスクで精度の向上が見込める.

本稿では, 金融ネットワークとりわけ銀行間取引ネットワークに着目し, それに適したデータ変換として softsign 関数を用いたデータ正規化手法や, 変分自己符号化器に基づく共変分自己符号化器 (Co-Variational Autoencoder: Co-VAE) ネットワーク埋め込みモデルを提案する. 提案モデルの性能評価は, ハイパーパラメータの探索の後, データの再構築と欠損したリンク属性の予測という観点から行う. また, Co-VAE を用いた銀行間取引における銀行間の金利予測に関する実験も行う.

2　関連研究

この節では, 提案モデルである Co-VAE に関連する自己符号化器 [1] と変分自己符号化器 [2] やネットワーク埋め込みに関して説明する.

2.1　自己符号化器

自己符号化器 (Autoencoder: AE) [1] は教師なし学習の一種であり, ニューラルネットワークを使用したデータの次元削減を行い, 潜在表現 (Latent representations)

*連絡先： 神戸大学 大学院 システム情報学研究科
〒 657-8501 兵庫県神戸市灘区六甲台町 1-1
E-mail: yuta@cs25.scitec.kobe-u.ac.jp
†連絡先： 広島大学 情報科学部
〒 739-8511 東広島市鏡山 1-3-2
E-mail: eguchi@acm.org

を獲得するモデルである. 自己符号化器の一番簡単なモデルは, 入力層, 潜在層, 出力層の 3 層から構成される.

このモデルの目標出力は入力自身であり, 入力を X とすると次のように表すことができる. 潜在層 Z は, $Z = f_\theta(X) = \phi(WX + b)$, 出力層 X' は, $X' = f'_{\theta'}(Z) = \phi'(W'Z + b')$ となる. ここで, W, W' は重みパラメータ, b, b' はバイアスパラメータであり, $\theta = (W, b)$, $\theta' = (W', b')$ とする. また, 関数 ϕ, ϕ' は活性化関数であり, シグモイド関数やランプ関数 (Rectified Linear Unit: ReLU) などが用いられる. 入力層の表現から潜在層の表現を得ることを符号化 (encode), 潜在層の表現から出力層の表現を得ることを復号化 (decode) と呼ぶ.

モデルパラメータ θ, θ' の学習は, 勾配降下法などを用いて損失関数 L を最小化することで行われる. ここで, 損失関数 $L(X, X')$ は入力層の表現 X と出力層の表現 X' の誤差で定義され, しばしば二乗誤差や交差エントロピーなどが用いられる.

2.2 変分自己符号化器

変分自己符号化器 (Variational Autoencoder: VAE) [2] は, 自己符号化器の一種であり, 深層生成モデルである. 自己符号化器と異なる点として, 潜在変数 $\boldsymbol{z}$ に確率分布を仮定していることがあげられる. 変分自己符号化器の目的はデータを生成する真の分布 $p_{\boldsymbol{\theta}^*}(\boldsymbol{x})$ を求めることであるが, その過程で入力データの潜在表現を獲得するパラメータを学習することができる.

独立同分布な N 個のサンプルからなるデータセット $\boldsymbol{X} = \{\boldsymbol{x}^{(i)}\}_{i=1}^N$ を考える. このデータが, 観測されない潜在変数 $\boldsymbol{z}$ を含む確率過程によって生成されるとすると, ある値 $\boldsymbol{z}^{(i)}$ がある事前分布 $p_{\boldsymbol{\theta}^*}(\boldsymbol{z})$ によって生成され, あるデータ $\boldsymbol{x}^{(i)}$ がある条件付き分布 $p_{\boldsymbol{\theta}^*}(\boldsymbol{x}|\boldsymbol{z})$ によって生成されると考えることができる. ここで $\boldsymbol{\theta}^*$ は真のパラメータである. しかし, 潜在変数 $\boldsymbol{z}^{(i)}$ と $\boldsymbol{\theta}^*$ は既知ではない. そこで, 複雑な真の事後分布 $p_{\boldsymbol{\theta}}(\boldsymbol{z}|\boldsymbol{x})$ を近似した符号化器 $q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x})$ を導入し, 復号化器 $p_{\boldsymbol{\theta}}(\boldsymbol{x}|\boldsymbol{z})$ とともにパラメータの学習を行う.

変分自己符号化器の学習は, 生成器 $p_{\boldsymbol{\theta}}(\boldsymbol{x})$ の対数尤度の最大化で行われる. しかし, 一般的に $p_{\boldsymbol{\theta}}(\boldsymbol{x})$ を直接計算することはできないため, 変分下限 (variational bound) の最大化を行う. 周辺尤度は $\log p_{\boldsymbol{\theta}}(\boldsymbol{X}) = \sum_{i=1}^N \log p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)})$ であり, 個々のデータ点 i での周辺尤度は次のようになる.

$$\log p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}) = D_{KL}(q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})||p_{\boldsymbol{\theta}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})) + \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}; \boldsymbol{x}^{(i)}).$$

右辺の第一項は真の事後分布とその近似との KL ダイバージェンス[2]である. KL ダイバージェンスとは, カルバック・ライブラー情報量とも呼ばれ, 2 つの確率分布の差異を測る尺度であり, 非負となる. また, 第二項は, データ点 i の周辺尤度の変分下限と呼ばれ, 次のように表される.

$$\begin{aligned}\log p_\theta(\boldsymbol{x}^{(i)}) &\geq \mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}; \boldsymbol{x}^{(i)}) \\ &= \mathbb{E}_{q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})}[-\log q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)}) + \log p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}, \boldsymbol{z})].\end{aligned}$$

この最右辺を変形すると式 (1) のようになり, これの最大化を行うことで個々のデータ点での周辺尤度が最大となり, 結果として, 周辺尤度 $\log p_{\boldsymbol{\theta}}(\boldsymbol{X})$ が最大化される.

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}; \boldsymbol{x}^{(i)}) &= -D_{KL}(q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})||p_{\boldsymbol{\theta}}(\boldsymbol{z})) \\ &\quad + \mathbb{E}_{q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})}[\log p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}|\boldsymbol{z})]. \quad (1)\end{aligned}$$

第一項に関して計算を行うために, $p_{\boldsymbol{\theta}}(\boldsymbol{z})$ に正規分布 $\mathcal{N}(\boldsymbol{z}; \boldsymbol{0}, I)$, $q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})$ に $\mathcal{N}(\boldsymbol{z}; \boldsymbol{\mu}, \boldsymbol{\sigma}^2)$ を仮定する. また, 損失関数を微分可能とするために, reparametrization trick [2] を用いて, $\boldsymbol{z}^{(i,l)} = \boldsymbol{\mu}^{(i)} + \boldsymbol{\sigma}^{(i)} \odot \boldsymbol{\epsilon}^{(l)}$ とする. ただし, $\odot$ はアダマール積すなわちベクトル要素ごとの積を意味し, $\boldsymbol{\epsilon}^{(l)}$ は $\boldsymbol{\epsilon}^{(l)} \sim \mathcal{N}(0, I)$ からの l 回目のサンプリングによって得ることを表す. 平均 $\boldsymbol{\mu}^{(i)}$ と標準偏差 $\boldsymbol{\sigma}^{(i)}$ は符号化器によって得られる. したがって, 第一項は次のようになる.

$$\begin{aligned}&-D_{KL}(q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})||p_{\boldsymbol{\theta}}(\boldsymbol{z})) = \\ &\quad \frac{1}{2}\sum_{j=1}^{J}(1 + \log((\boldsymbol{\sigma}_j^{(i)})^2) - (\boldsymbol{\mu}_j^{(i)})^2 - (\boldsymbol{\sigma}_j^{(i)})^2). \quad (2)\end{aligned}$$

式 (1) の第二項は, あるデータ $\boldsymbol{x}_i$ が与えられたとき, $q_{\boldsymbol{\phi}}(\boldsymbol{z}|\boldsymbol{x}^{(i)})$ から $\boldsymbol{z}$ をサンプリングし, その後, $p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}|\boldsymbol{z})$ からサンプリングして得られた $\boldsymbol{x}^{(i)}$ が, 入力として与えた $\boldsymbol{x}^{(i)}$ である度合いを表している. すなわち, どれだけうまく再構築できたかを示しており, 再構築誤差とも呼ばれる. $p_{\boldsymbol{\theta}}(\boldsymbol{x}^{(i)}|\boldsymbol{z})$ からサンプリングして得られた $\boldsymbol{x}^{(i)}$ を $\boldsymbol{x}'^{(i)}$ と表記したとき, 再構築誤差に二乗誤差を用いると誤差の最小化, すなわち, $-||\boldsymbol{x}^{(i)} - \boldsymbol{x}'^{(i)}||_F^2$ の最大化と考えることもできる. 以上をまとめると式 (3) のようになる.

$$\begin{aligned}\mathcal{L}(\boldsymbol{\theta}, \boldsymbol{\phi}; \boldsymbol{x}^{(i)}) &= \frac{1}{2}\sum_{j=1}^{J}(1 + \log((\boldsymbol{\sigma}_j^{(i)})^2) - (\boldsymbol{\mu}_j^{(i)})^2 - (\boldsymbol{\sigma}_j^{(i)})^2) \\ &\quad -||\boldsymbol{x}^{(i)} - \boldsymbol{x}'^{(i)}||_F^2. \quad (3)\end{aligned}$$

ここで, J は潜在表現の次元数, $||\cdot||_F$ はフロベニウスノルムである. なお, 行列 A の i 行 j 列の要素を a_{ij} と表記すると A のフロベニウスノルムは $||A||_F = \sqrt{\sum_i \sum_j a_{ij}^2}$ と定義される.

2.3 ネットワーク埋め込み

ネットワーク埋め込み (グラフ埋め込み) は, ネットワークデータのノードやリンクなどに関して分析する

[2] P と Q を連続確率分布とするとき, $D_{KL}(P||Q) = \int P(x) \log \frac{P(x)}{Q(x)} dx$ と定義される.

ことで何らかの洞察を得ることを目的とする．一般的に, ネットワーク埋め込みはネットワーク構造の低次元空間への写像と捉えることができ, その写像によって保存されるネットワーク構造はネットワーク埋め込み手法に依存する．ネットワーク埋め込みには様々な手法が存在しており, Goyal らは文献 [3] で因子分解ベースの手法, ランダムウォークベースの手法, 深層学習ベースの手法とその他に分類している．また, それらの手法に関して一般的なデータセットを用いた性能比較を行っている．

ネットワークの埋め込み手法で考慮すべき重要な観点として, ネットワーク構造の保存があげられる．その中でも, 1 次近接性と 2 次近接性といったノード間の類似性の尺度を表す近接性という概念がある．1 次近接性とは, ノード v_i とノード v_j 間の隣接性またはリンクの重みに着目するものであり, 局所的な構造を表す．2 次近接性とは, ノード v_i とノード v_j の近傍構造の類似性に着目するものであり, より大域的な構造を表す．2 次近接性の構造を保存することで近傍構造の類似性からリンクの欠損を捉えやすくなると考えられる．そのため, リンク予測などのタスクでは 2 次近接性を保持した埋め込みを行うことがより重要となる．

本稿で提案する Co-VAE によるネットワーク埋め込みは, 2 次近接性を保持する埋め込み手法に位置づけられる．

3 提案手法

ネットワークは, リンクの向きの有無により, 無向グラフと有向グラフに分類することができ, 本稿の実験で対象とする銀行間取引データは重み付き有向グラフで表現することができる．この有向グラフからは, 各ノードに入ってくるリンク (インリンク) に着目するか, 各ノードから出ていくリンク (アウトリンク) に着目するかで 2 通りの隣接行列を得ることができる．本稿で提案するのは, このリンクの向きを考慮した上で, それぞれの潜在表現を同時に学習するモデルであり, 共変分自己符号化器 (Co-Variational Autoencoder: Co-VAE) と名付ける．この節では, 無向グラフにおける VAE に基づくネットワーク埋め込みモデルに関する説明の後, 提案モデルである Co-VAE の概要を示し, 有向グラフにおける VAE や Co-VAE に基づくネットワーク埋め込みモデルに関して述べる．

3.1 無向グラフにおける VAE に基づくネットワーク埋め込み

ノード数が N である重み付き無向グラフのネットワークを考える．このネットワークのノード間の隣接関係, すなわちリンクの有無とリンクの重みは隣接行列で表現することができ, これは対称行列となる．あるノード v_i の隣接表現を表すベクトル $s_i = [s_{i,0}, s_{i,1}, \ldots, s_{i,n-1}]$ が入力データとなる．ただし, ノード v_i とノード v_j が隣接しない場合は, 対応する成分が零となり, 隣接する場合は, v_i と v_j の間のリンクの重みをとる．この隣接表現を入力とするモデルの学習により, 非線形な写像による再構築を行う．

本稿での基礎技術となる変分自己符号化器のモデルの構造を図 1 に示す．符号化器 (encoder) は, 入力層と 2 層の中間層, そして, 2 層目の中間層の出力と正規分布 $N(0,1)$ から ϵ をサンプリングすることで得られる潜在表現を出力とする層から構成される．また, 復号化器 (decoder) は, 潜在表現を入力とする層と 2 層の中間層, 再構築された表現を出力とする出力層から構成される．

このモデルのパラメータ学習は, 式 (3) に基づいた損失関数 (式 (4)) を最小化することで得られる．パラメータの最適化手法には Adam [4] などを用いる．

$$loss = \frac{1}{N}\sum_{i=1}^{N}\Big\{p_\alpha \cdot \Big(-\frac{1}{2}\Big)\cdot\sum_{j=1}^{J}\Big(1+\log\big((\sigma_j^{(i)})^2\big)-(\mu_j^{(i)})^2 - (\sigma_j^{(i)})^2\Big) + (1-p_\alpha)\cdot||p_\beta(x^{(i)}-x'^{(i)})||_F^2\Big\}. \quad (4)$$

第一項は, ネットワーク埋め込みで得られる潜在表現が正規分布とどれだけ近いかを表しており, 第二項は, 入力と出力の二乗誤差を表す．ここで, N はデータ数, J は潜在表現 z の次元数, p_α は第一項と第二項を調整するハイパーパラメータであり, $x'^{(i)}$ は学習する VAE に $x^{(i)}$ を入力したときの出力である．また, p_β は零要素と比較して非零要素の誤差をどれだけ重要視するかの程度を表すパラメータである．

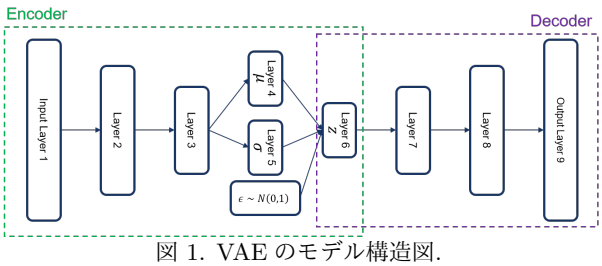


図 1. VAE のモデル構造図.

3.2 有向グラフにおける VAE に基づくネットワーク埋め込み

次に, ノード数が N である重み付き有向グラフのネットワークを考える．このネットワークでは, 各ノードに入ってくるリンク (インリンク) に着目するか, 各ノードから出ていくリンク (アウトリンク) に着目するかで 2 通りの $N \times N$ の隣接行列を得ることができる．ただし, 有向グラフの隣接行列は一般に非対称行列となることに注意が必要である．すなわち, インリンクに着目した隣接行列を S とすると, アウトリンクに着目した

隣接行列は S の転置行列である S^T で表され, 一般に $S \neq S^T$ となる. このうち, どちらをモデルの入力データとするか, すなわち, どちらの行列における行ベクトルに着目し, モデルを用いた次元削減を行うかによって, 得られる潜在表現と解釈が異なってくる[1]. このことから, 本稿では変分自己符号化器 (VAE) ネットワーク埋め込みモデルとそれを拡張した共変分自己符号化器 (Co-VAE) ネットワーク埋め込みモデルを提案する. VAE ネットワーク埋め込みモデルはインリンクに着目した潜在表現を獲得するモデル, もしくはアウトリンクに着目した潜在表現を獲得するモデルであり, モデルの構造としては, 無向グラフと同様に図 1 で表される. Co-VAE ネットワーク埋め込みモデルはインリンクに着目した潜在表現を獲得する符号化器とアウトリンクに着目した潜在表現を獲得する符号化器を同時に学習し, 両者の潜在表現から復号化を行うモデルである. このモデルの構造は, 図 2 で表される. VAE ネットワーク埋め込みモデルと Co-VAE ネットワーク埋め込みモデルは, ともに式 (4) を最小化することでパラメータを学習する.

3.3 時系列を考慮した学習

ネットワークが時系列的に変化するとき, その変化を捉えつつネットワーク埋め込みを行うことが重要となる. そこで, 前述の VAE と Co-VAE に基づいたネットワーク埋め込みモデルにおける時系列を考慮した学習として次の 3 種類を提案する.

1 つ目が, モデルパラメータ継承法である. これは, ある時刻 t で学習したモデルの重みパラメータのすべてを時刻 $t+1$ のモデルの重みパラメータの初期値として継承し, 学習を行う方法である. すなわち, ネットワークの時系列的な変化の大きさに応じてパラメータ更新を行う手法であると解釈できる. ある時刻 t と時刻 $t+1$ でネットワークに変化がないときは時刻 t の重みパラメータを維持し, ネットワークの変化が大きいときは時刻 t の重みパラメータの影響を小さくしつつ, 時刻 $t+1$ のデータから得られる特性を強く反映するパラメータ更新を行う. この手法では, 損失関数に変更はなく, 式 (4) の最小化によりパラメータを学習する.

2 つ目が, 深層部パラメータ制約法である. これは, 前節までで述べた VAE ネットワーク埋め込みモデルおよび Co-VAE ネットワーク埋め込みモデルの符号化器の 2 つ目の中間層から μ, σ の層へと順方向に伝播することを表す重みパラメータを前の時刻のパラメータと近くなるように制約を課して学習を行う方法である. μ と σ につながる重みパラメータは, モデルの構造上一番次元数が削減されている部分であるため, 深層部パラメータ制約法と呼ぶことにする. この手法では, 式 (4) に, 該当する重みパラメータの誤差項を追加した式 (5) を最小化することでパラメータを学習する.

$$loss = \frac{1}{N}\sum_{i=1}^{N}\Big\{p_\alpha\left(-\frac{1}{2}\right)\sum_{j=1}^{J}\Big(1+\log((\sigma_j^{(i)})^2)-(\mu_j^{(i)})^2-(\sigma_j^{(i)})^2\Big) + (1-p_\alpha)||p_\beta(x^{(i)}-x'^{(i)})||_F^2 + p_\gamma||W_{d_t}-W_{d_{t-1}}||_F^2\Big\}. \quad (5)$$

ここで, p_γ はその影響力を調整するハイパーパラメータであり, W_{d_t} と $W_{d_{t-1}}$ は学習中の t 時刻のモデルと, 前時刻の $t-1$ の学習済みモデルの該当箇所の重みパラメータである.

3 つ目は, モデルパラメータ制約法と深層部パラメータ制約法を組み合わせる方法である.

4 実験データ

実験に用いたデータは欧州債務危機と呼ばれる時期を含む, 2009 年 7 月 1 日から 2012 年 12 月 31 日におけるイタリアの銀行を主とした, 欧州の銀行間取引のデータである (https://www.e-mid.it/en/). このデータは, 取引の契約日, 取引を持ちかけられた側の銀行（Quoter）, 取引を持ちかけた側の銀行（Agressor）, 取引金額, 取引の金利, 取引の状態などの要素から構成されている. この期間の内, 取引に関わった銀行は 153 行, 取引件数は 162,075 件であった. 取引金額は百万ユーロ単位で示されており, 取引の状態は”Sell”と”Buy”の 2 種類がある.

取引状態”Sell”とは, Agressor が Quoter に”Sell”という取引を持ちかけたものである. すなわち, 取引状態”Sell”では, Agressor から Quoter に取引金額分の金額が動く. 取引状態”Buy”とは, Agressor が Quoter に”Buy”という取引を持ちかけたものである. すなわち, 取引状態”Buy”では, Quoter から Agressor に取引金額分の金額が動く. この 2 種類の取引状態をもとに, 一定期間で集計することで重み付き有向グラフを作成することができる. 本稿では, 月単位で集計した重み付き有向グラフを表す, 42ヶ月分の隣接行列 (153×153) を時系列性のあるデータセットとして実験を行う. 以下, 時刻 t の隣接行列を S_t と表記し, i 行 j 列の要素を $S_{t_{ij}}$ とする. ただし, $t \in \{1, \ldots, 42\}$ である. この隣接行列の要素 $S_{t_{ij}}$ は, 銀行 i から銀行 j に $S_{t_{ij}}$ 分の金額が動いたことを表している. すなわち, この隣接行列の行ベクトルは銀行 i が他の銀行にお金を貸したという関係 (アウトリンク) に着目した表現である. また, この隣接行列を転置することで得られる S_t^T は行ベクトルが他の銀行からお金を借りたという関係 (インリンク) に着目した表現である. この節以降, それぞれの表現を貸し手側, 借り手側と呼ぶ.

データ S_t は, 23409($= 153 \times 153$) の要素で構成されている. このデータセットは非常にスパースであり, 要

[1]無向グラフでは隣接行列は対称行列となるため, $S = S^T$ となり, どちらの行列からでも得られる潜在表現と解釈は同じである.

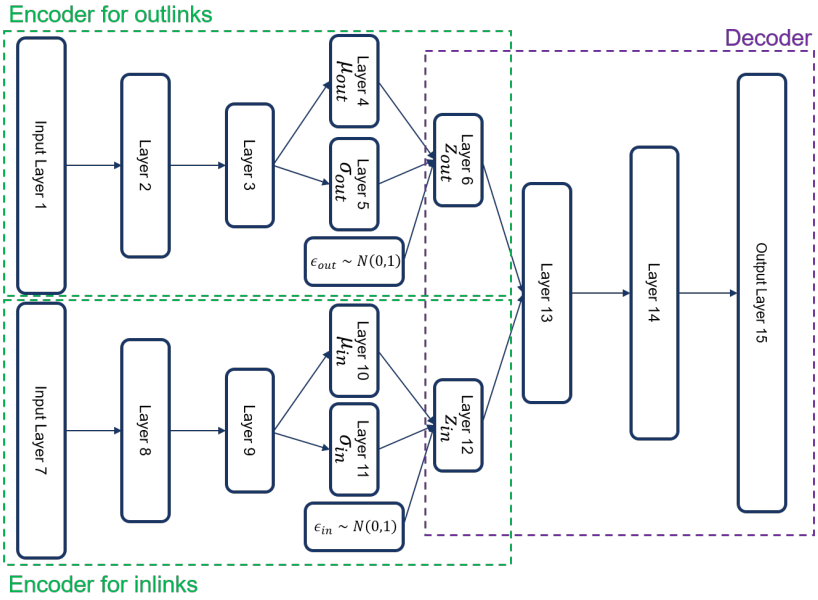


図 2. Co-VAE のモデル構造図.

素の値の度数分布をヒストグラムで表すと 0 に偏った分布となる. また, 非零要素の平均値はおよそ 95.7, 中央値は 25.0 となる. 銀行間の取引時に決定される金利についても同様に隣接行列が作成できる. ただし, この隣接行列の要素の値は, 総和ではなく期間内にノード間で行われた取引での金利の平均値とする. なお, この隣接行列は R_t で表す.

4.1 データセットの前処理

ニューラルネットワークを用いた機械学習を行う際には, 適切に正規化などのデータ変換を学習前に行うことがしばしば求められる. 代表的なデータ変換手法として, min-max 正規化などがあげられる. min-max 正規化とは, 与えられたデータ X の最小値と最大値を用いて, 最小値を 0, 最大値を 1 に線形変換するようなデータ変換である. これは, $\tilde{x}_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}$ により得られる. ここで, x_{min} と x_{max} はデータ X の最小値と最大値である.

min-max 正規化には外れ値の影響をうけやすいことや, 最小値や最大値が既知であることを前提としているなどの問題点が存在していた. そこで, 本稿では, 実データの分布を意識した非線形なデータ変換として, softsign 関数を用いた正規化を提案する. softsign は文献 [5] で, 活性化関数として扱われており, tanh と同程度の性能を示すと述べられている. なお, 文献 [5] では softsign 関数を活性化関数として用いたのに対し, 本稿ではデータ変換として用いる.

softsign 正規化 softsign 関数は $\mathrm{softsign}(x_i) = \frac{x_i}{1+|x_i|}$ で表される. この関数の特徴として, $\mathrm{softsign}(x) \in (-1, 1)$ や $x_i = 1$ で $\tilde{x}_i = 0.5$ となることなどがあげられる. この特徴を活かしつつ, 定数 1 をパラメータ α で置き換えることで変換後の分布を制御することが可能となる. これを式で表すと $\mathrm{softsign}_\alpha(x_i) = \frac{x_i}{\alpha+|x_i|}$ となり, これを用いた正規化手法を $\mathrm{softsign}_\alpha$ 正規化とする. 様々なパラメータ α における $\mathrm{softsign}_\alpha$ の関数の形状を図 3 に示す. この正規化手法の特徴としては, 最大値と最小値が既知でなくてよいことや, 非線形な変換であるがデータの大小関係は保持されること, 分布の偏りをパラメータ α によって制御できることなどがあげられる.

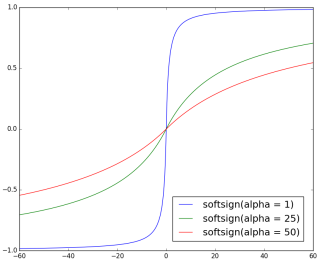


図 3. softsign 関数の波形 ($\alpha = 1, 25$).

実際に, min-max 正規化と $\mathrm{softsign}_\alpha$ 正規化を適用したデータセット全体での要素の値の度数分布をヒストグラムで表して, 図 4 に示す. ただし, 縦軸は対数スケールとしており, ビンの数は 50 である. min-max 正規化による変換では, ヒストグラムの概形はデータ変換前の概形と変わらず, 0 に偏った分布となる. また, $\mathrm{softsign}_1$ 正規化は, 非零要素と零要素を分割した概形となる. $\mathrm{softsign}_{25}$ 正規化は, データセットの非零要素

の中央値である 25.0 で 0.5 を取るようになっており, 非零要素の数がグラフにおいて左右に概ね均等に分割される.

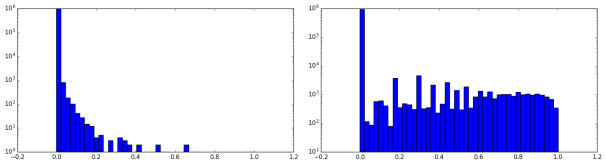


図 4. データ変換後のヒストグラム. (左)min-max. (右)softsign_25.

4.2 データセットの分割

6 節で述べる実験は検証実験とテスト実験からなり, 検証実験でハイパーパラメータの探索を行い, そこで決定されたハイパーパラメータを用いてテスト実験を行う. そのため, データセットを時系列性を考慮した上で検証実験用データセットとテスト実験用データセットに分割する. 検証実験に用いる学習データを, $S_t(1 \leq t \leq 36)$ とし, 検証データとしては S_{t+1} を用いる. すなわち, 学習済みモデルに検証データを入力し, 出力をもとに評価値を算出する. また, テスト実験に用いる学習データを, $S_t(37 \leq t \leq 41)$ とし, テストデータとして S_{t+1} を用いる. すなわち, 学習済みモデルにテストデータを入力し, 出力をもとに評価値を算出する. 次の節で述べるリンク予測評価は, 実際に取引が行われている非零のリンクのうち, 15%のリンクをランダムに隠したデータを用いて評価値を算出する. この 15%を隠したデータセットは, S_t 毎に 10 通り用意している.

5 評価方法

この節では, 提案モデルの評価方法である再構築評価とリンク予測評価や各評価指標について説明をする.

再構築評価 本稿では再構築評価として, 入力データ[2]から符号化・復号化を経て入力データを再現した出力がどれだけランキングを保持しているか, すなわちデータ値の大きさに基づく順序関係を保持して再現されているかを測る. ランキングの一貫性を評価するために, 後述のケンドール順位相関係数を用いる. 変分自己符号化器の特徴として, 同一の入力データであっても, 符号化器の出力を計算する際に, 正規分布 $\mathcal{N}(0,1)$ から ϵ のサンプリングを行っているため, 若干のズレが生じる. そのため, 入力データと出力データの相関を複数回測り, その平均値をもって再構築の評価に用いる.

リンク予測評価 リンク予測評価とは, 一般的に学習データに存在しなかったリンクの有無を予測するものであるが, 本稿では, 学習データに存在しなかったリンクの重みを予測することを指す. そのため, 4.2 節でも述べたように, 予め観測されているデータの一部を隠すことで未観測のリンクを作成し, その欠損データ (欠損値を含むデータ) を学習済みのモデルに入力する. これにより, 欠損データの補完がどれだけランキングを保持して行えているかを測る. 再構築評価と同様に, ランキングの保持の評価には後述のケンドールの順位相関係数を用いる. 未観測のリンクとして予め隠すリンクの重要度により, 補完の精度に影響を及ぼすことが考えられる. したがって, 予め用意した共通の欠損データを用いて, モデルの評価毎に欠損部の差異がないようにする. また, 欠損部により精度が変わることが考えられるため, 複数の欠損データで評価を行い, その平均をもってリンク予測評価の評価値とする.

ケンドール順位相関係数 ケンドール順位相関係数 τ (Kendall's tau) は 2 つのランキングデータの相関性を測るノンパラメトリックな評価尺度である. これは, τ で表記され, -1 から 1 までの範囲の値を取る. $|\tau|$ が大きいほど強い相関があることを示す.

平均二乗誤差 平均二乗誤差 (Mean Squared Error: MSE) は, 真値と予測値の誤差を測る尺度である. データ数が N, 真値が $y_i \in Y$, 予測値が $\hat{y}_i \in \hat{Y}$ であるとき, $MSE(Y,\hat{Y}) = \frac{1}{N}\sum_{i=0}^{N-1}(y_i - \hat{y}_i)^2$ と定義される.

平均絶対パーセント誤差 平均絶対パーセント誤差 (Mean Absolute Percentage Error: MAPE) は真値に対する予測値の誤差をパーセンテージで表すものである. データ数が N, 真値が $y_i \in Y$, 予測値が $\hat{y}_i \in \hat{Y}$ であるとき, $MAPE(Y,\hat{Y}) = \frac{100}{N}\sum_{i=0}^{N-1}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$ と定義される.

6 実験

本節では, 提案モデルの性能を確認するための実験を行う. また, 4 節で述べたように, 銀行間の取引ネットワークのデータは, 貸し手側に着目した隣接行列と借り手側に着目した隣接行列を獲得することができ, 貸し手側の隣接行列の行ベクトルは融資状況の表現, 借り手側の隣接行列の行ベクトルは借金状況の表現といえる. したがって, 融資状況の表現から VAE ネットワーク埋め込みモデルを用いて融資パターンの潜在表現を獲得することを貸し手着目モデル, 借金状況の表現から VAE ネットワーク埋め込みモデルを用いて借金パターンの潜在表現を獲得することを借り手着目モデル, Co-VAE ネットワーク埋め込みモデルで融資パターンと借金パターンの両者の潜在表現を同時に獲得することを貸し手借り手同時着目モデルと呼ぶこととする.

本実験では, 正規化手法や出力層の活性化関数, 時系列考慮手法などの組み合わせをいくつか用いた. よって, 表 1 のように正規化手法や出力層の活性化関数の

[2]本稿における実験では, 4 節で定義した隣接行列 S, またはそれに対して 4 節に示した正規化を施した行列を指す.

組み合わせを Config_ID で略記する．また, 時系列考慮手法として, モデルパラメータ継承法, 深層部パラメータ制約法, 両者の組み合わせの 3 種類で実験を行うが, それぞれ, Config_ID の後部に _v1, _v2, _v3 を加えることで表記する．

表 1: 各 Config ID が示す正規化手法と出力層の活性化関数の組み合わせ．

Config IDs	正規化手法	出力層の活性化関数
Config_01	min-max	linear
Config_02	softsign_1	linear
Config_03	softsign_1	sigmoid
Config_04	softsign_{25}	sigmoid
Config_05	softsign_{25}	relu

各 Config_ID で, 貸し手着目モデル, 借り手着目モデル, 貸し手借り手同時着目モデルで重みパラメータの学習を行う．このとき, 各 Config_ID に対する着目モデルごとに, 損失関数である式 (4) や式 (5) のハイパーパラメータの格子探索を行う．格子探索の範囲は, $p_\alpha = [0.9, 0.99, 0.999]$, $p_\beta = [1, 10, 20, 40, 80]$, $p_\gamma = [1, 10, 100, 1000, 10000]$ とする．また, 貸し手着目モデルと借り手同時着目モデルの各層の次元数は図 1 を参考にすると, Layer 番号順に (153, 100, 50, 10, 10, 10, 50, 100, 153) であり, 貸し手借り手着目モデルの各層の次元数は図 2 を参考にすると, Layer 番号順に, (153, 100, 50, 10, 10, 10, 153, 100, 50, 10, 10, 10, 100, 200, 306) である．パラメータ更新は Adam を用いており, 学習率は 0.001, エポック数は 1000 とし, 出力層以外の活性化関数には relu を用いている．これらを通して, 各 Config_ID の着目モデルごとに最適パラメータを検証データを用いて決定し, テストデータを用いて性能を比較する．

まずはじめに, 表 1 の各設定で着目モデルごとに学習を行い, 再構築評価とリンク予測評価の 2 つの観点から比較する．ここでは, 貸し手借り手同時着目モデルの結果のみを表 2 と表 3 に示すが, 他の 2 つでも同様の傾向となった．結果として, min-max 正規化よりも softsign 正規化がデータ変換として有効であることが示せた．また, 両評価タスクで Config_05 が良好な評価値を示すことが確認できた．さらに, 評価タスクによって最適なハイパーパラメータが異なることも確認できた．

再構築評価では, データセット全体のランキングの相関を測っているため, 零や非零を区別しつつ再構築を行っていることが求められる．その結果, 誤差項を重要視する p_α がより大きな値をとるようになったと考えられる．一方, リンク予測評価では, 隠されたリンクを再構築した結果どれだけ大小関係を捉えているかを測っている．したがって, 隠されていない部分から構造的特徴を捉えた再構築を行うことが求められる．その結果, 非零要素を重要視する p_β がより大きな値をとるようになったと考えられる．これ以降は, 両評価タスクで良い性能を示した Config_05 の設定でのリンク予測評価に関する結果を示す．

次に, Config_05 の設定における各着目モデルの性能比較を行い, Co-VAE ネットワーク埋め込みモデルが優れた性能を示すかを確認する．VAE ネットワーク埋め込みモデルでは捉えられていなかった情報を貸し手側と借り手側の潜在表現を同時に学習する Co-VAE ネットワーク埋め込みモデルが捉えることで, 優れた性能を示すと期待される．表 4 に示した結果において, Co-VAE を用いた貸し手借り手同時着目モデルは他よりも優れた評価値を示した．

表 2: 再構築評価における貸し手借り手同時着目モデルの各設定での最適ハイパーパラメータと Kendall's tau.

Config IDs	p_α	p_β	学習データ	検証データ
Config_01	0.900	1	0.182 ± 0.015	0.166 ± 0.015
Config_02	0.999	40	0.279 ± 0.025	0.251 ± 0.024
Config_03	0.999	10	0.283 ± 0.026	0.252 ± 0.024
Config_04	0.999	10	0.282 ± 0.026	0.252 ± 0.024
Config_05	0.999	10	0.588 ± 0.019	0.461 ± 0.027

表 3: リンク予測評価における貸し手借り手着目モデルの各設定での最適ハイパーパラメータと Kendall's tau.

Config IDs	p_α	p_β	検証データ
Config_01	0.900	80	0.278 ± 0.025
Config_02	0.900	80	0.278 ± 0.029
Config_03	0.990	80	0.299 ± 0.027
Config_04	0.990	80	0.347 ± 0.034
Config_05	0.990	80	0.335 ± 0.034

表 4: リンク予測評価における各着目モデルでの最適ハイパーパラメータと Kendall's tau.

Models	p_α	p_β	検証データ	テストデータ
貸し手着目	0.9	80	0.312 ± 0.045	0.335 ± 0.034
借り手着目	0.9	80	0.327 ± 0.039	0.342 ± 0.059
同時着目	0.99	80	0.335 ± 0.034	**0.375 ± 0.026**

次に, 時系列を考慮した学習法としてモデルパラメータ継承法 (v1), 深層部パラメータ制約法 (v2), 両者の組み合わせ (v3) の比較実験を Config_05 の設定で行う．この結果を貸し手借り手着目モデルのみ表 5 に示す．深層部パラメータ制約法 ($v2$) が最良であった．時系列を考慮することで評価値自体はやや低下しているものの, 時区間を横断した分析が可能となるのが利点である．

最後に, 得られた潜在表現を用いた応用として, ニューラルネットワークによる回帰モデルを用いた金利予測の実験を行う．これは, 入力層, 中間層, 出力層の 3 層で構成される．入力はネットワーク埋め込みによって得ら

表 5: リンク予測評価での各時系列考慮法に対する貸し手借り手同時着目モデルの最適ハイパーパラメータと Kendall's tau.

Config IDs	p_α	p_β	p_γ	検証データ	テストデータ
Config_05_v1	0.999	40	-	0.290 ± 0.028	0.265 ± 0.032
Config_05_v2	0.999	20	1000	0.325 ± 0.042	**0.345 ± 0.025**
Config_05_v3	0.990	10	100	0.296 ± 0.045	0.322 ± 0.024

れた潜在表現を利用することで $X =$ [時刻 t の i 銀行の融資パターンの潜在表現, 時刻 t の j 銀行の借金パターンの潜在表現] と連接し, 目標出力を $Y = R_{t_{i,j}}$ とする回帰モデルの学習を行う. ここで, 活性化関数は relu, パラメータ更新には Adam を用いて学習率は 0.003 とした. また, 学習や評価に用いた $R_{t_{i,j}}$ は, t 月の金利の隣接行列の要素のうち非零である, すなわち取引を行っているものを対象とした. この $R_{t_{i,j}}$ は, i 銀行が貸し手側, j 銀行が借り手側のときに設定される金利を意味している.

VAE ネットワーク埋め込みモデルは, 貸し手着目モデルと借り手着目モデルの 2 つを用いてそれぞれの潜在表現を獲得することで X を作成する. Co-VAE ネットワーク埋め込みモデルは, 貸し手側の潜在表現と借り手側の潜在表現を同時に獲得できるため, これのみで X を作成する. このとき, X は 20 次元であり, 中間層は 5 次元とした. また, それぞれに深層部パラメータ制約法を適用したときに得られる潜在表現を用いた予測も行う. なお, これらの着目モデルのハイパーパラメータはリンク予測評価で最適となったパラメータセットを用いる. 比較対象として, softsign_{25} 正規化を適用後の元の隣接行列を用いた予測も行う. このとき, 入力データは $X =$ [i 銀行の貸し手側の隣接表現, j 銀行の借り手側の隣接表現] の 306 次元からなり, 中間層の次元数は 18 に設定した. これらの予測性能は, MAPE で評価する.

それぞれの金利予測の結果, 元の隣接行列を用いたときは 58.82%, Config_05 の VAE を用いたときは 55.41%, Co-VAE を用いたときは 50.16%, Config_05_v2 の VAE を用いたときは 52.13%, Co-VAE を用いたときは 49.75% となった. 結果として, Co-VAE を用いることや深層部パラメータ制約法を適用することによる性能の向上が確認できた.

7 むすび

本稿では, 銀行間取引ネットワークの分析を行う手法として, 共変分自己符号化器 (Co-VAE) ネットワーク埋め込みモデルを提案した. これは, ノード間のインリンクとアウトリンクに着目するモデルであり, 実験ではインリンクに着目した借金パターンの潜在表現とアウトリンクに着目した融資パターンの潜在表現を同時に学習した. また, softsign_α 正規化や時系列を考慮した学習法も提案した.

実験の結果, softsign_α 正規化はモデルの性能向上に大きく貢献したといえる. また, 応用上重要とされるリンク予測評価においては提案モデルである Co-VAE ネットワーク埋め込みモデルの有効性と時系列を考慮した学習として深層部パラメータ制約法が有効であることが確認できた. 金利予測に関する実験でも良好な結果を示した.

今後の課題としては, より詳細なハイパーパラメータ探索をともなう精緻な分析・評価があげられる. 本稿では, Co-VAE の復号化器における中間層の次元数を VAE ネットワーク埋め込みモデルの 2 倍と設定したが, 適切な次元数を選択することで, さらなる精度の向上が見込める.

謝辞

本研究を行うにあたり，有益な助言を頂いた神戸大学大学院経済学研究科の羽森茂之教授と金京拓司教授に感謝する．本研究に使用したデータを提供して頂いた神戸大学大学院システム情報学研究科谷口隆晴准教授に感謝する．本研究の一部は科学研究費補助金基盤研究 (B)（15H02703）の援助による.

参考文献

[1] Geoffrey Hinton and Ruslan Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, Vol.313, No.5786, pp.504-507, (2006)

[2] Diederik P. Kingma and Max Welling, "Auto-Encoding Variational Bayes," *ICLR 2014*, (2014)

[3] Palash Goyal and Emilio Ferrara, "Graph Embedding Techniques, Applications, and Performance: A Survey.," *CoRR*, (2017)

[4] Diederik P. Kingma and Jimmy Ba, "Adam: A Method for Stochastic Optimization," *ICLR 2015*, (2015)

[5] Xavier Glorot and Yoshua Bengio, "Understanding the difficulty of training deep feedforward neural networks," *AISTATS'10*, (2010)

自己符号化器を用いた 国債イールドカーブのファクターモデルの構築

Autoencoder-Based Three-Factor Model for Yield Curve of Japanese Government Bond

水門善之 [1,2*] 坂地泰紀 [1]
和泉潔 [1] 島田尚 [1] 松島裕康 [1]

Yoshiyuki Suimon[1,2] Hiroki Sakaji[1]
Kiyoshi Izumi[1] Takashi Shimada[1] Hiroyasu Matsushima[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: 本研究では，機械学習手法の一種である自己符号化器（オートエンコーダ）を用いて，日本国債のイールドカーブの 3 ファクターモデルの構築を行った．更に，構築した自己符号化器を形成するニューラルネットワークのモデルパラメータに着目することで，自動生成した 3 つのファクターが，それぞれ，イールドカーブの水準・曲率・傾きを表現していることを示した．加えて，本研究では，構築した自己符号化器を，国債の割高・割安の判別器として使用することで，国債のロングショート戦略を構築した場合，トレンドフォロー型の投資戦略に比べて，良好なパフォーマンスが得られることを確認した．

1. はじめに

金利とは，人々の経済活動の活況度合いを映し出す代表的な指標である．また，国内におけるマクロ経済環境や様々な政策要因のみならず，海外市場における金利変動要因が，様々な金利裁定取引を介して国内の金利市場に伝達されるメカニズムも存在している．国債の年限別の金利をつなぎ合わせたイールドカーブは，これら様々な要因を反映しながら日々変動する．本研究では，機械学習手法の一種である自己符号化器（オートエンコーダ）を用いて，日本国債のイールドカーブを表現する 3 ファクターモデルの構築を行う．更に，構築した自己符号化器を形成するニューラルネットワークのモデルパラメータに着目することで，自動生成した 3 つのファクターが，それぞれイールドカーブの水準・曲率・傾きを表現していることを確認する． 加えて，構築した自己符号化器を，国債の割高・割安の判別器として使用する投資戦略についても検証を行う．

2. イールドカーブの期間構造モデル

2.1 金利の変動

金利（国債の利回り）は，マクロ経済環境を表す代表的指標である．一般に，経済活動が活発な時は，様々な資金需要の高まりを背景に，銀行貸出や社債，国債等の金利に上昇圧力がかかる．また，国債や社債等の金利は発行体である国や企業のクレジットリスクの影響を受ける．他にも，海外市場における金利変動要因も，様々な金融裁定取引を通じて国内市場に影響を与える．このように，多様な市場環境の変化を織り込みながら，国債の市場価格は形成され，そこから計算される金利（国債の利回り）も日々変化していく．

* Email: d2018ysuimon@socsim.org

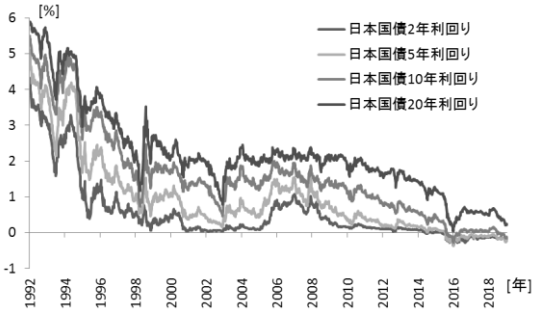


図 1: 日本国債利回りの推移

2.2 国債イールドカーブの期間構造モデル

図 1 に示した通り，金利には様々な年限のものが存在する．現在，日本では短期の割引国債に加え，満期が 2 年，5 年，10 年，20 年，30 年，40 年の利付国債などが発行されており，幅広い年限の市場金利が存在している．このような異なる年限の金利をつなぎ合わせた曲線はイールドカーブと呼ばれる（図 2）．

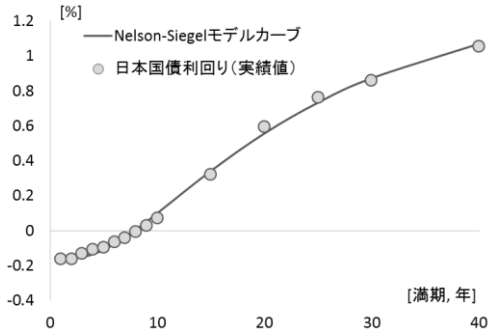


図 2: Nelson-Siegel モデルに基づくイールドカーブ

イールドカーブの形状は，金利の期間構造モデルを用いて表すことができる. 例えば, Nelson-Siegel モデル[1]を用いると，以下のような関数形を用いてイールドカーブの期間構造が表現できる．ここでは y は金利，τ は年限，λ は定数とする．

$$y(\tau) = F_1 + F_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + F_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \quad (1)$$

モデルの関数は，イールドカーブの水準，傾き，曲率を表す 3 つのファクターF の項から成り立っており，それぞれのファクターの年限別の係数は図 3 のようになる．ただし，Nelson-Siegel モデルは，イールドカーブの形状を簡便な関数によって近似的に表現するというモデルであることから，利便性が高い反面，関数型の制約を受けやすいといった側面もある．

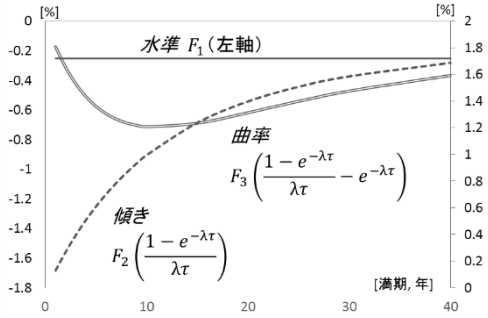


図 3: Nelson-Siegel モデルの各ファクターの係数

一方で，金利データに対して直接主成分分析を行うことで，主要なファクターを抽出する方法も知られている．1992 年以降の 2，5，7，10，15，20 年の日本国債利回りの週次の水準データを対象に，主成分分析を行うと，第 3 ファクターまでの累積寄与度は約 99%となり，概ねイールドカーブの形状が表現できることが分かる．更に，図 4 に各ファクターの固有ベクトルの形状を掲載した．これによると，第 1，第 2，第 3 ファクターが，それぞれ，イールドカーブの水準，傾き，曲率を意味することが分かる．

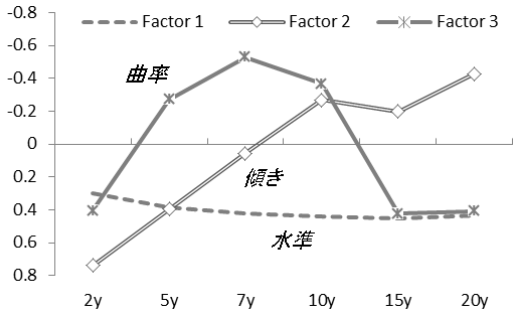


図 4: 主成分分析に基づく固有ベクトル

また，金利の期間構造モデルの構築に機械学習手法を用いたものとしては，水準・傾き・曲率の情報を軸に，ニューラルネットワークを用いてイールドカーブの変動をモデル化した研究[2]が挙げられる．更に，これを拡張させる形で，日米イールドカーブの金利情報を直接ニューラルネットワークモデルに組み込んだ研究[3][4]や，金利平価説を踏まえて日米金利に加えてドル円レートも同一のニューラルネットワークに組み込んだ研究[5]も挙げられる．

3. 自己符号化器を用いたモデル構築

3.1 自己符号化器

自己符号化器（オートエンコーダ，Autoencoder）とは，機械学習手法の一種であり，ニューラルネットワークを使用した次元圧縮のためのアルゴリズム

として知られる[6]．線形の次元圧縮としては主成分分析が挙げられるが，自己符号化器とはニューラルネットワークの入力層と出力層に，同じデータを教師データとして学習させる手法であり，本研究では3層のニューラルネットワークモデルを構築する．

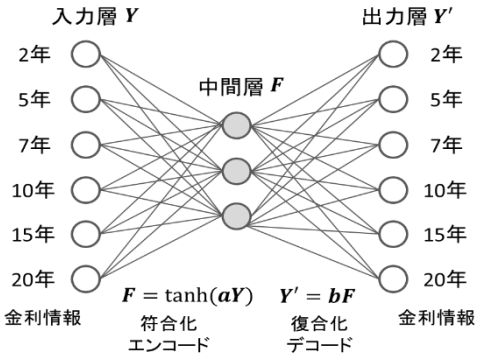


図 5: 自己符号化器

本研究では，図5に示すように2,5,7,10,15,20年の金利情報を用いた学習を行う．学習には1992年7月から2019年7月までの週次データを用いる．更に，出力層へのパスを線形の関数とし，その係数に着目することで，自己符号化器における中間層の各ノードの解釈を行う．

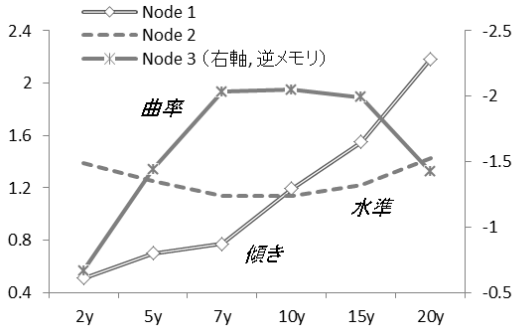


図 6: 中間層から出力する線形関数の係数

図6には，中間層からの出力を表す線形関数について，出力層のノード別（各年限別）の係数を示した．これによると，中間層の各ノードが，イールドカーブにおける水準，傾き，曲率を意味していると解釈できよう．これらを踏まえ，各ノードの変化を，実際の金利の水準，金利差（スプレッド）と比較したものが図7である．例えば，水準を表すと解釈されるノード2を，イールドカーブの起点に近い2年金利と比べると両者の概ねの連動が見られる．また，傾きを表すと解釈できるノード1と，2-20年の金利スプレッド（＝20年利回り－2年利回り）を比較すると，両者は概ね近い動きをしていることが分かる．更に，図6によるとノード3は長期金利を中心としたイールドカーブの曲率を表すと解釈できるが，実際，2-10-20年のバタフライスプレッド（＝2×10年利回り－10年利回り－20年利回り）と比較すると，両者が概ね近い動きをしていることが確認できる．

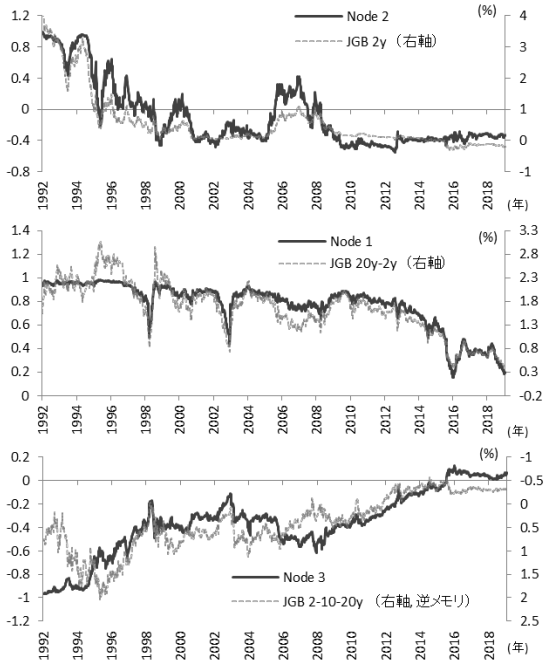


図 7: 中間層のノードと金利水準及び金利差

3.2 金利予測シミュレーション

本節では，学習済みの自己符号化器が出力する金利が，イールドカーブ上の他年限の金利との過去の相対関係に基づいて算出される点を踏まえ，学習した自己符号化器を，国債の割高・割安の判別器として用いる国債のロングショート戦略を構築する．

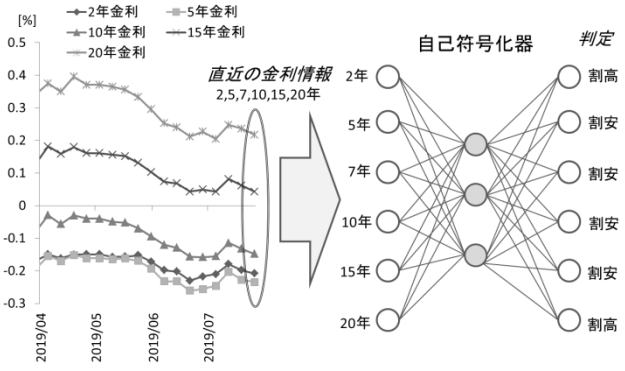


図 8: 自己符号化器を用いた割高・割安の判定

具体的には，各時点における各年限の金利情報を，学習済みの自己符号化器に入力し，そこから出力される金利を基準金利として，実際の金利が基準金利よりも高い場合は，割安と判断して当該年限の国債

をロング（買い）し，実際の金利が基準金利よりも低い場合は割高と判断しショート（売り）する．いずれのポジションも保有期間は 1 カ月及び 3 カ月とする．本シミュレーションの結果は図 9 に示した．ここでは，各ロングショート戦略の 1 カ月間の平均キャピタルゲインを掲載している．単位は bp（=0.01%）であり，例えばロングの場合は，投資期間中の金利低下幅，ショートの場合は金利上昇幅をキャピタルゲインとする．本分析では，金利予測精度の検証のため，キャリー及びロールダウン，またショートの際のレポコストの影響は考慮しない．自己符号化器の学習には，モデル更新時点のデータを除く後方 5 年間のサンプルデータを使用し，モデルは 1 年ごとに更新を行う．自己符号化器の学習期間を，後方 10 年間とした場合の結果も図 10 に掲載した．これらの検証には 1992 年 7 月から 2019 年 7 月までの週次の金利データを用いた．

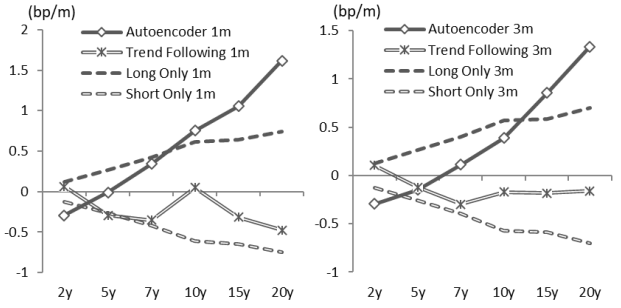


図 9: 1 カ月間の平均キャピタルゲイン
（モデル学習期間 5 年，保有期間 1 カ月・3 カ月）

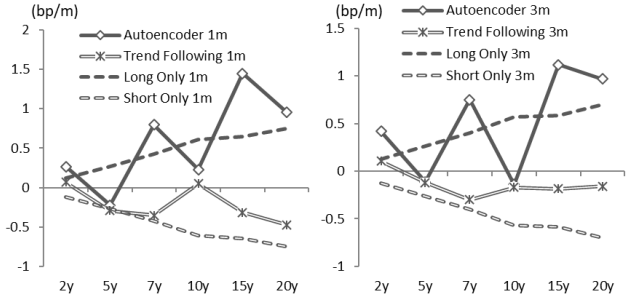


図 10: 1 カ月間の平均キャピタルゲイン
（モデル学習期間 10 年，保有期間 1 カ月・3 カ月）

比較対象として，トレンドフォロー型の投資戦略（各投資時点で前週から金利が低下していればロング，上昇していればはショート），及び常にロング（ショート）を続けた戦略の結果も併せて掲載した．これによると，本提案モデルはトレンドフォロー型の投資戦略に比べて，幅広い年限で良好なパフォーマンスが得られることが確認できる．本モデルは，学習済みの自己符号化器が，他年限金利との相対関係を基に，基準金利を算出することから，仮に対象年限の金利が他年限対比で歪んでいた場合，その歪みの修正に向けたポジションが自動構築できる点が特長である．

4. 結論

本研究では，自己符号化器（オートエンコーダ）を用いて，日本国債のイールドカーブの 3 ファクターモデルの構築を行った．更に，ニューラルネットワークのモデルパラメータに着目することで，中間層の 3 つのファクターが，それぞれ，イールドカーブの水準・曲率・傾きを表現していることを示した．加えて，本研究では構築した自己符号化器を，国債の割高・割安の判別器として使用することで，国債のロングショート戦略を構築した場合，トレンドフォロー型の投資戦略に比べて，良好なパフォーマンスが得られることを確認した．

参考文献

[1] Charles R Nelson and Andrew F Siegel, “Parsimonious Modeling of Yield Curves,” The Journal of Business, Vol.60, No.4, pp.473-489, 1987
[2] 水門善之，“機械学習を用いた国債イールドカーブの変動モデルの構築と長期金利予測,” 人工知能学会第 21 回金融情報学研究会, No.21, pp.46-49, 2018
[3] 水門善之，坂地泰紀，和泉潔，島田尚，松島裕康,“日米イールドカーブの連動性を用いた機械学習に基づく日本国債の長期金利予測,” 人工知能学会第 22 回金融情報学研究会, pp.81-87, 2019
[4] Y. Suimon, H. Sakaji, K. Izumi, T. Shimada, H. Matsushima, “Japanese long-term interest rate forecast considering the connection between the Japanese and US yield curve,” IEEE Conference on Computational Intelligence for Financial Engineering and Economics, 2019
[5] Y. Suimon, H. Sakaji, K. Izumi, T. Shimada, H. Matsushima, “Extraction of relationship between Japanese and US interest rates using machine learning methods,” International Conference on Smart Computing and Artificial Intelligence, 2019
[6] G. E. Hinton and R. R. Salakhutdinov, “Reducing the Dimensionality of Data with Neural Networks,” Science, Vol.313, No.5786, pp.504-507, 2006

留意事項

本稿は，著者の個人見解を表すものであり，野村證券株式会社の公式見解を表すものではありません．

日銀「主な意見」の発言者分類モデルの作成と政策変更予測への応用

Speaker Classification Model on BOJ's Summary of Opinions for Forecasting Their Policy Change

末廣徹 [1*,2]　木村柚里 [1]　稲垣真太郎 [1]

Toru Suehiro[1,2], Yuri Kimura[1], and Shintaro Inagaki [1]

[1] みずほ証券金融市場調査部
[1] Fixed Income Research Department, Mizuho Securities, Co., Ltd.
[2] 法政大学大学院経済学研究科経済学専攻
[2] Graduate School of Economics, Hosei University

Abstract: In this paper, we introduce several types of text classification model that predict the speakers on BOJ's "Summary of Opinions". Those documents summarize BOJ's Monetary Policy Meetings and the major part of the contents consist of the committee members' opinions, but those comments are kept anonymous. Regarding this issue, the commentator prediction model should be a great help for focusing the next decision of BOJ. Our models are trained with the past public speech texts of BOJ committees. In order to correct the bias of datasets, we tried some data pre-processing before model fitting. The proposal models are Random Forest, LSTM and Bidirectional LSTM with attention mechanism. As a result, we achieved over 90% accuracy for the best. To applying those models to the analysis of BOJ's "Summary of Opinions", we focus on the relationship between the ratio of presumed speakers in the documents and past monetary policy changes. The analysis revealed that the higher ratio of "Reflationist's" assertion raises the chance of monetary easing occurrence.

1. はじめに

13 年 4 月の異次元緩和政策の導入以降，日銀は日本国債の買い入れを急増し，19 年 3 月末時点では市中残高の約 43%を保有するようになった．しかし，安定的な物価目標 2%の達成は困難な状況が続いており，更なる追加緩和策導入の可能性も十分にある．他方，中長期的には超低金利環境における金融機関の運用難と収益悪化の帰結であるリバーサルレートの考え方から，日銀は早期に金融政策を正常化するとの見方も根強い．いずれにせよ，債券市場を中心とした市場参加者にとって日銀の次の一手を予想することは経済環境を予想することよりも重要なタスクになりつつある．いわゆる「BOJ ウォッチ」（中銀ウォッチ）と言われる日銀（その他各国中銀）の分析は，中央銀行が発表する金融政策決定会合の声明文や議事要旨，各委員の個別の講演内容などを精読することで，ニュアンスの変化を捉えるという方法が一般的である．証券会社やシンクタンクに所属する民間エコノミスト・アナリスト（BOJ ウォッチャー）が独特な言い回し（日銀文学とも言われる）を読み込み，ちょっとした変化を読み取る．たとえば，16 年 1 月のマイナス金利導入の決定などを事前に予想することができていれば，金融市場においてそれ相応に収益を上げることができただろう．

日銀の金融政策決定会合は，総裁と 2 人の副総裁に 6 人の審議委員を加えた 9 人の政策委員による合議制で決定される（多数決）．そのため，政策の決定内容だけでなく，各委員が独自に行う講演や匿名で公表される議事要旨などのテキストを読み込む必要がある．匿名で公表される「主な意見」や「議事要旨」といった資料において，それぞれの発言が誰のものであるかを正確に予測することができれば，決定会合における議論の方向性もつかみやすくなる．

* 連絡先: みずほ証券株式会社金融市場調査部
〒100-0004 東京都千代田区大手町 1-5-1 大手町ファーストスクエア
E-mail: toru.suehiro@mizuho-sc.com

本稿では，このような大量のテキストの特徴理解をテキストマイニング（文字情報のデータ化）と機械学習・深層学習の組み合わせによって行った．そして，筆者らが作成したモデルを用いて新規のテキストの発言者の予想（分類）を行い，決定会合における匿名の発言内容が政策委員の誰の意見によっているのかを検証し，決定会合における各委員の発言比率と 16 年以降に断続的に行われてきた金融緩和策導入のタイミングとの関連性を調べた．

2. 先行研究

日銀が公表するテキストやデータを機械学習させて分析したものには，日銀の金融経済月報を共起解析，主成分分析，回帰分析の三段階からなる CPR 法を用いてテキスト解析することで，市場価格の変動を予想した研究として[1]がある．[1]は，日経平均，円ドル，国債(2,5,10 年)利回りの変動をそれぞれ一定の精度で予測できることを示した(例えば，日経平均については 70%超)．

また，日銀の金融政策に関する研究としては，[2]がある．これは日銀の政策決定変更のパターンを，日銀が公表する決定会合の声明文，金融経済月報，展望レポート，総裁記者会見録などの文書を深層学習によって抽出し，政策変更の予想に役立てる研究を行った．具体的には，Doc2Vec によって獲得した文章の分散表現と時系列データと組み合わせて深層信念ネットワーク（DBN）に学習させる，新たに入力された日銀の文書が，次回の決定会合で金融政策の変更を行った過去ケースの特徴を有するか否かを判定する分類器を開発した．分類器は過去 2 回（2014 年 10 月と 2016 年 1 月）の金融政策変更に対して，ある程度有用なシグナルを発したことを示した．

そのほか，[3]は，日銀会合後の総裁記者会見における，黒田総裁の表情と金融政策変更の関係を研究した．[3]は深層学習等を用いた表情認識アルゴリズムを用いることで，会見における総裁の表情を解析し，重大な金融政策変更を行う直前の回の会見では，「怒り」や「嫌悪」の表情が多くなる一方，政策変更後の会見では「悲しみ」の数値が低下する傾向を確認した．

これらの研究は日銀の公表するテキスト等の情報を機械学習や深層学習によって分析し，市場価格や金融政策変更の予想を行ったものだが，本稿が注目する 9 名の日銀政策委員による個々の発言を分析したものとしては[4]がある．[4]は 9 名の政策委員の講演テキスト（17 年 1 月～18 年 9 月）について，Random Forest を用いることで各委員の発言が誰のものかという正答率は約 52%となることを示した．本稿は[4]の分析を拡張させたものである．

3. データとモデル

本稿では，日銀政策委員の講演テキストを機械学習モデルによって学習した「日銀政策委員の発言分類器」を作成し，予測精度を検証する．また，作成した分類器を用いて，匿名で委員の発言が公表されている日銀金融政策決定会合「主な意見」の発言者予測を行い，各決定会合でどの委員の発言比率が高いかという予測結果と，金融政策決定会合における政策変更の関係を考察する．

3.1 日銀政策委員の講演

日銀政策委員は各種講演やイベントにおいて挨拶を行うことがある．日銀 HP にアップされているだけでも 18 年に 37 回の講演・挨拶を行っており，黒田総裁が 19 回，雨宮副総裁が 4 回となっている．その他審議委員は地方における「金融経済懇談会」で挨拶をすることが多く，それぞれの年間平均で約 2 回行っている．これらの講演・挨拶では，政策委員会全体の「公式見解」を説明しつつも，各委員が特に注目している内容や，独自の考え方を説明するケースがある．講演テキストを学習データとすることで，日銀政策委員の発言を分類するモデル（分類器）を作成することができる．

3.2 「主な意見」とは

作成した分類器を用いる対象（ターゲットデータ）として，日銀金融政策決定会合「主な意見」を用いる．「主な意見」とは日銀の金融政策決定会合の 6 営業日後をめどに公表される資料で，9 人の政策委員が会合で発言した内容について，整理されて公表される．日銀によると「金融政策決定会合における主な意見」は，(1) 各政策委員および政府出席者が，金融政策決定会合で表明した意見について，発言者自身で一定の文字数以内に要約し，議長である総裁に提出する，(2) 議長はこれを自身の責任において項目ごとに編集する，というプロセスで作成されたものである．

「主な意見」のコメントは様々な委員からある程度は均等に選ばれているとみられるが，金融・経済の環境によっては発言比率に偏りがみられ，金融政策変更との関連性があると予想される．

3.3 分析データと前処理

日銀政策委員の発言に関する「分類器」を作成するための入力データは 2016 年 1 月～2020 年 1 月までの日銀政策委員の講演テキストとした．学習データの数は全体で 16,605 センテンスあり，黒田総裁がそのうち 8,689 センテンスと圧倒的に多く，次いで，原田審議委員が 2,456 センテンス，最も少なかったのは片岡審議委員の 347 センテンスであった．

文書分類モデル構築にあたって，テキストデータにいくつかの処理を施した．まず，講演テキストは「はじめに」の節で「最近の金融・経済の環境」を中心とした一般的な説明を行うケースが多く，各委員の特徴が現れにくい．また，「おわりに」の節も同様で，講演が行われた地方の経済の特色を述べるなど，委員の基本的な金融経済に対する主張を分類する際のノイズとなり得るため，各委員の講演テキストからこれらを削除した．その結果，センテンス数は全体で 14,583 センテンスとなった．

さらに，黒田総裁の発言数の多さを解消するため，黒田総裁の発言文書の数をランダムに選び出し 2,303 センテンスまで削減し，2 番目にセンテンス数が多い原田審議委員とのオーダーを合わせた．その結果，全体で 9,150 センテンスとなった．

それでもなおカテゴリー（9 人の委員）におけるセンテンス数の偏りが残るため，文書数の少ない各委員（黒田総裁と原田審議委員以外の 7 人分）については，文書拡張処理によってセンテンスのオーダーを揃えた．本研究では[5]によって検証された Easy Data Augmentation(EDA)の手法に従った文書拡張を行った．すなわち，データサイズ 9,150 に対し，日本語 Wordnet シソーラスによる類義語置換(Synonym Replacement)を行った．[5]の研究結果を参考に，各文書に対して同義文を生成し，データセットに加えた．その結果，各委員のセンテンス数は 2,000～2,800 となり，全体では 21,184 センテンスとなった．

表 1: データサイズ一覧

	Row	Crop1	Crop2	Crop+aug
黒田	8,689	7,736	2,303	2,303
雨宮	994	822	822	2,415
若田部	534	462	462	2,138
原田	2,456	2,254	2,254	2,254
布野	1,099	905	905	2,635
櫻井	1,113	946	946	2,778
政井	481	395	395	2,266
鈴木	892	789	789	2,323
片岡	347	274	274	2,072
計	**16,605**	**14,583**	**9,150**	**21,184**

注: 「row」は元データ，「crop1」は「はじめに」と「おわりに」を削除したもの，「crop2」は黒田総裁のセンテンスを一部削減したもの，「aug」は文書拡張後のデータ．

3.4 COS 類似度による比較

機械学習モデルの作成に先立ち，各委員の講演テキストの COS 類似度を比較した．COS 類似度を比較することにより，どの委員がより特徴的な発言をしやすいかどうかの検証が可能になる．具体的には各委員の講演テキストおよび 9 人すべての講演テキストを Bag of Words によってベクトル化し，各委員の相対での COS 類似度を求めた．類似度が高ければ，他の委員と類似した発言が多いことになる．他方，他の委員との類似度が低ければ，独自の発言が多く，特徴が得やすい委員である可能性が高い．

3.5 機械学習モデル

前述の前処理(3.3)を行った各委員の講演テキストに対して，機械学習モデルを作成した．本稿では，モデルの精度を比較検討するため，3 つのモデルを作成した．具体的には，いずれも前述した前処理を行った上で，①Random Forest，②LSTM，③Attention 機構を用いた Bidirectional LSTM の 3 つである．

これら 3 つを使用することでモデルによる精度の違いを検証する．本稿では，LSTM に加えて順方向の伝播だけでなく逆方向の伝播も考慮された Bidirectional LSTM も用いた．この理由としては[6]において，Bidirectional LSTM は複数判別のケースにおける文書分類を行う場合 LSTM よりも精度が高いことが示されたからである．さらに，[7]によると，Bidirectional LSTM に Attention 機構を適用することで，精度の向上がみられるという結果があることから，これらを参考にモデルの選定を行った．

3.6 金融緩和策導入との関係

16 年以降の「主な意見」について，どの委員の発言に近いものが多かったかを調べ，実際の金融政策の変更との関係を分析する．具体的には 16 年以降の「主な意見」について，各委員の発言比率と，過去の「金融緩和イベント」（0 と 1 のダミー変数）との関係を，2 項ロジスティック回帰分析を用いて調べた．「金融緩和イベント」は「マイナス金利導入の決定」（16 年 1 月）や「ETF 買入の増額」（17 年 7 月）や「物価目標の達成時期の延期（廃止）」（16 年 4・11 月，17 年 7 月，18 年 4 月），「フォワードガイダンスの明確化」（19 年 4 月）を「金融緩和イベント」とした．

4. 分析結果と考察

4.1 COS 類似度による比較の結果

各委員の講演テキストを Bag of Words によってベクトル化し，各委員の相対での COS 類似度を求めた（表 2）．

類似の平均値が最も高いのは黒田総裁だった．これは黒田総裁は日銀の代表として，様々な委員の考えを述べることが多いことなどが理由とみられる．

表 2: 各政策委員の発言の COS 類似度

	黒田	雨宮	若田部	原田	布野	櫻井	政井	鈴木	片岡
黒田	1.00	0.96	0.95	0.91	0.91	0.94	0.96	0.93	0.94
雨宮		1.00	0.91	0.88	0.86	0.90	0.92	0.89	0.91
若田部			1.00	0.88	0.90	0.93	0.94	0.91	0.90
原田				1.00	0.81	0.85	0.86	0.83	0.85
布野					1.00	0.95	0.94	0.94	0.90
櫻井						1.00	0.95	0.93	0.93
政井							1.00	0.93	0.94
鈴木								1.00	0.91
片岡									1.00

4.2 機械学習モデルの精度

本節では，①Random Forest，②LSTM，③Attention 機構を用いた Bidirectional LSTM の 3 つのモデルについて，精度の検証を行う．

まず①Random Forest について，学習データ（前処理済みの講演テキスト）からランダムに選んだ 80%を学習データ，残りの 20%をテストデータとして正答率を複数回検証した結果（Cross Validation）は 72.5%となった．次に，②LSTM，③Attention 機構を用いた Bidirectional LSTM についても正答率を求めた．それぞれの正答率は 89.6%，92.0%となり，正答率は改善した．LSTM を用いた方が Random Forest よりも正答率が上がることや，[6]と同様に，Bidirectional LSTM によってさらに精度が向上することが示された（表 3，4）．

表 3: 各モデルの正答率（%）

Model	Accuracy
① RandomForest	72.5
② LSTM	89.6
③ BiLSTM+Att	92.0

表 4: 各発言者に対する正答率（%）

	LSTM	BLSTM＋Att
黒田	69.0	75.8
雨宮	89.3	93.5
若田部	91.1	92.5
原田	89.0	92.3
布野	92.7	92.0
櫻井	92.5	93.8
政井	87.3	92.9
鈴木	96.3	96.3
片岡	99.3	98.2

なお，Biderectional LSTM+Att のモデルについて，それぞれの委員の発言の正答率を個別に求めると（Confusion Matrix），表 5 のようになった．比較的 COS 類似度が高かった黒田総裁と若田部副総裁の分類を間違える比率が高いなどの特徴が得られた，

表 5: BLSTM+Att による各委員の発言の正答率(%)

		正解								
		黒田	雨宮	若田部	原田	布野	櫻井	政井	鈴木	片岡
予測	黒田	69.3	3.2	6.9	5.1	4.0	4.2	1.5	5.7	0.2
	雨宮	1.8	94.9	0.7	0.7	0.0	1.4	0.0	0.5	0.0
	若田部	14.9	2.7	79.0	1.2	0.5	1.7	0.0	0.0	0.0
	原田	7.7	0.0	0.0	87.6	0.6	0.9	0.9	2.4	0.0
	布野	2.3	0.0	0.2	1.0	94.1	0.2	0.2	2.0	0.0
	櫻井	3.6	1.4	0.5	1.4	0.7	90.2	0.0	2.0	0.2
	政井	1.0	0.0	0.0	0.0	1.0	0.2	97.7	0.0	0.0
	鈴木	1.3	0.8	0.0	0.8	0.8	0.2	0.8	94.9	0.2
	片岡	0.0	0.5	0.0	0.0	0.2	0.2	0.5	0.9	97.7

4.3 金融緩和策との関係

前節で示した 3 つのモデルのうち，②LSTM のモデルを用いて，16 年以降の毎回の「主な意見」の発言比率を分析した．具体的には，発言者が不明の「主な意見」の発言集に対して，モデルを適用し，それぞれの発言がどの委員の発言に近いかを検証した．毎回の「主な意見」において，誰の発言が多かったのかを調べた結果は図 1 に示した通りである．前述したように，本来は各委員が均等に選ばれている可能性が高いものの，推計された発言比率は会合によ

って発言者（の予測）に偏りが生じた．

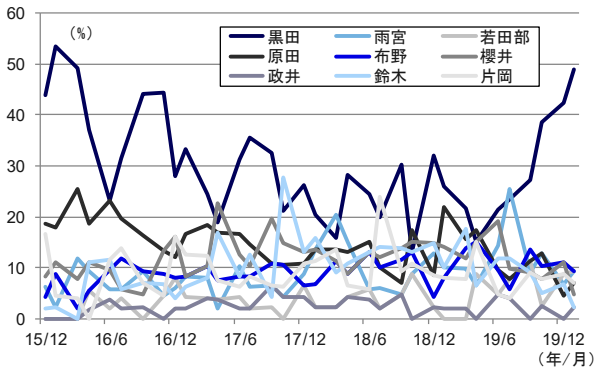


図 1: 「主な意見」のコメントを分類器モデル（BLSTM+Att）によって分類した場合の比率

委員の発言比率と金融緩和策との関係を調べるため，「金融緩和イベント」ありの決定会合を「1」とし，「政策変更なし」の決定会合を「0」としたダミー変数を作成し，それぞれの委員の発言比率を説明変数とした 2 項ロジスティック回帰分析を行った（表 6）．

回帰係数の P 値をみると，10%の有意水準で，金融緩和イベントダミーに対して原田審議委員の発言比率の回帰係数がプラス（発言比率が高いと金融緩和が行われる可能性が高い）となることが分かった．統計的に有意とは言えないまでも，原田審議委員の発言比率の説明力が高いことが分かる．金融緩和が行われる決定会合では，積極的な金融緩和を主張することの多いリフレ派である原田審議委員に近い発言が増えることが予想される．また，原田審議委員と同じリフレ派とされる片岡審議委員の発言比率も加えた「リフレ派」（原田＋片岡）の発言比率の回帰結果をみると，5%の有意水準で回帰係数がプラスとなることが分かる．つまり，現在の日銀の政策委員において，リフレ派の委員が金融緩和政策の導入を促してきた可能性が高い．今後の金融政策決定会合においても，リフレ派の発言比率が高くなることがあれば，追加緩和策導入の機運が高まる可能性がある．

反対に 5%の有意水準で、金融緩和イベントダミーに対して政井審議委員と鈴木審議委員の発言比率の回帰係数がマイナス（発言比率が高いと金融緩和が行われる可能性が低い）となることが分かった．

表 6: 「金融緩和イベント」のダミー変数を被説明変数とし，各委員の「主な意見」における発言比率（モデルによる推測）を説明変数とした 2 項ロジスティック回帰分析の結果

	係数	標準偏差	z値	P<|z|	#	Pseudo R^2	Log likelihood
黒田	5.8	5.1	1.1	0.25	34	0.05	-11.65
雨宮	-42.6	24.4	-1.8	0.08	34	0.20	-9.91
若田部	-10.7	18.6	-0.6	0.57	34	0.01	-12.14
原田	22.1	13.2	1.7	0.09	34	0.14	-10.62
布野	10.7	19.3	0.6	0.58	34	0.01	-12.16
櫻井	-14.4	15.3	-0.9	0.35	34	0.04	-11.83
政井	-113.8	56.4	-2.0	0.04	34	0.28	-8.88
鈴木	-37.1	18.4	-2.0	0.04	34	0.27	-8.97
片岡	15.8	10.8	1.5	0.14	34	0.09	-11.22
原田＋片岡	25.7	12.8	2.0	0.05	34	0.26	-9.07

5. まとめと今後の課題

本稿の分析結果により，日銀の政策委員の発言を分類するモデルを作成し，金融政策の変更（金融緩和策の導入）の予測に用いることについて，一定の成果が得られることが分かった．

日銀の政策委員の発言を分類するモデルの作成については，①Random Forest，②LSTM，③Attention 機構を用いた Bidirectional LSTM の 3 つのモデルの精度を検証し，それぞれのモデルの特徴を分析した．

これらの比較により，モデルの精度については Random Forest＜LSTM＜Bidirectional LSTM となることを示した．Bidirectional LSTM を用いたモデルが最も精度が高いという結果は先行研究である[6]と同じである．

最後に，委員の発言比率と金融緩和策との関係を調べるため，「金融緩和イベント」ありの決定会合を「1」とし，「政策変更なし」の決定会合を「0」としたダミー変数を作成し，それぞれの委員の発言比率を説明変数とした 2 項ロジスティック回帰分析を行った結果，委員によっては発言比率が金融政策の変更と一定の関係を有することが分かった．特に，金融緩和イベントに対して，金融緩和を主張することが多い「リフレ派」と呼ばれる原田委員と片岡委員の発言比率の関係性が最も深いことは，「リフレ派」の主張内容とも整合的である．データの蓄積を進め，一段の検証を行いたい．

参考文献

[1] 和泉 潔, 後藤 卓, 松井 藤五郎: テキスト分析による金融取引の実評価, 第 24 回人工知能学会全国大会論文集,3H1-OS12a-2,(2010)

[2] 塩野 剛志: 文書の分散表現と深層学習を用いた日銀

政策変更の予想,第 16 回人工知能学会金融情報研究会資料 SI-FIN-016-11, (2016)

[３] 水門 善之, 勇 大地: 日銀総裁会見の表情解析に基づく感情値の計測と金融政策変更との関係, 第 19 回人工知能学会金融情報研究会資料 SI-FIN-019-22, (2017)

[４] 末廣 徹: AI・機械学習は「BOJ ウォッチ」できるのか〜 「主な意見」を分類してみた, みずほ証券金融市場調査部 Macro Information (2018 年 9 月 28 日), (2018)

[５] Wei, J. W., and Zou, K.: Eda: Easy data augmentation techniques for boosting performance on text classification tasks, arXiv preprint , arXiv:1901.11196, (2019)

[６] 山本 裕樹, 松尾 豊: 景気ウォッチャー調査の深層学習を用いた金融レポートの指数化, 第 30 回人工知能学会(2016)

[７] Zhou, P., Shi, W., Tian, J., Qi, Z., Li, B., Hao, H., and Xu, B.: Attention-based bidirectional long short-term memory networks for relation classification, Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Volume 2: Short Papers, pp.207-212, (2016)

ニュース記事に基づく景気指標S-APIRの開発

関 和広[1*]　生田 祐介[2]　松林 洋一[3,4]
Kazuhiro Seki[1]　Yusuke Ikuta[2]　Yoichi Matsubayashi[3,4]

[1] 甲南大学 知能情報学部
Faculty of Intelligence and Informatics, Konan University
[2] 大阪産業大学 経営学部
Faculty of Business Management, Osaka Sangyo University
[3] 神戸大学大学院 経済学研究科
Graduate School of Economics, Kobe University
[4] アジア太平洋研究所
Asia Pacific Institute of Research

Abstract: This paper reports our work on developing a new business sentiment index using daily newspaper articles. We adopts a recurrent neural network (RNN) to estimate the sentiment of a given text. The RNN is initially trained on Economy Watchers Survey and then fine-tuned on newspaper articles for domain adaptation. Also, a one-class support vector machine is applied to filter out texts of irrelevant topics. Moreover, we analyze the contributions of various factors that influence the estimated business sentiment.

1　はじめに

政府や中央銀行が定期的に発表する景気ウォッチャー調査[1]の現状・先行き判断 DI（diffusion index; 景気動向指数）や日銀短観[2]の DI などの経済指数は，金融当局の政策決定や企業の生産計画，機関投資家・個人の投資判断など，様々な経済活動の拠りどころとして重要な役割を担っている．しかしながら，これらの指数算出の基礎となる調査には大変な手間とコストがかかる．

たとえば，景気ウォッチャー調査の現状・先行き判断 DI（以降「景気ウォッチャー DI」と呼ぶ）の場合，北海道，東北，北関東，南関東，甲信越，東海，北陸，近畿，中国，四国，九州，沖縄の 12 地域を対象に，地域の景気に関連の深い動きを観察できる立場にある人々（小売店の店主，タクシー運転手など）2,050 人を対象に地域別調査機関がアンケート調査を行い，取りまとめ調査機関において地域ごとの調査結果を集計・分析し，指数の算出を行っている．また，このような調査には相応の時間を要するため，日銀短観は四半期に一度，景気ウォッチャー指数は一ヶ月に一度しか発表されず，リアルタイム性が低い．

一方，店舗の商品売り上げのような統計データ，各種ニュース記事やマイクロブログなどのテキストデータ，検索エンジンのクエリログデータ，スマートフォン等から収集される GPS の位置情報データ，あるいは人工衛星の衛星画像データなど，いわゆるオルタナティブデータが世界中で絶えず生成・収集され，大量に蓄積されている．このようなデータが利用可能になってきたこと，および深層学習（deep learning）に代表されるデータドリブンな人工知能技術の発展から，人手による調査を要さず，すでに存在する他の大規模なデータを代替的に用いた経済・金融指数の将来予測（forecast），あるいは足元予測（nowcast）への注目が高まっている [Kapetanios 18, 関 17]． 具体的な取り組みとして，たとえば店舗の POS（point of sale）データを用いた消費者物価指数のナウキャスト [Watanabe 14]，金融・経済レポートからの景況感指数の予測 [山本 16]，新聞記事からの穀物の市場価格や経済指数の予測 [Chakraborty 16], [Shapiro 17] などがある．

本研究では特にテキスト情報に着目し，日々発信されるニュース記事を利用して，新しい景況感指数（S-APIR 指数と呼ぶ）を提案する．また，予測された指数を用いて，特定の概念（語句）が景況感にどのように影響を与えたのか時系列で分析する．

*連絡先：甲南大学知能情報学部知能情報学科
〒 658-8501 兵庫県神戸市東灘区岡本 8-9-1
E-mail: seki@konan-u.ac.jp
[1]http://www5.cao.go.jp/keizai3/watcher/watcher_menu.html
[2]http://www.boj.or.jp/statistics/tk/

表 1: 景気ウォッチャー調査（現状の景気）の例.

地域	回答者	景気	判断理由
北海道	タクシー運転手	×	売上が減少しているが、季節要因に加えて、景気が下向いていることも影響している。
北関東	輸送用機械器具製造業	◎	自動車は米国への輸出が増えている。

2 関連研究

経済・金融分野では，失業率や物価指数などの統計値や個別企業の売上高など多くの数値データに加えて，新聞記事や決算短信，金融レポート等，関連するテキストデータが豊富に存在する．通常，これらのテキスト情報は人が読み込み，他の情報・状況と考えあわせ，総合的な意思決定を行うために利用される．しかしながら，すべての関連するテキスト情報を限られた時間で咀嚼・勘案することは，熟練した専門家であっても現実的には困難である [関 17]．そこで，このようなテキスト情報を用いて経済・金融指数を予測する試みが広く行われている．ここでは，本研究に関連が深い景況感指数（ビジネスセンチメント）の予測について代表的な研究をいくつか紹介する．

2.1 景況感指数の予測

1 節で述べた内閣府の景気ウォッチャー調査では，集計された指数（景気ウォッチャー DI）に加えて，個々のアンケート回答者が判断した現状および先行きの景気の状況（「◎（良い）」から「×（悪い）」までの 5 段階評価）とその理由（自然言語文）が公開されている．いくつかの例を抜粋して表 1 に示す．

景気ウォッチャー DI は，表 1 の「景気」の 5 段階の判断にそれぞれ点数を与え，これらの点数を各判断区分の構成比（%）に乗じた和で算出されている．算出された DI は，50 を基準（景気に変化なし）として 0 から 100 の値を取る．なお，景気ウォッチャー調査では，景気の現状判断と先行き判断を区別して回答させるため，それぞれに指数が算出される．

山本ら [山本 16] は，この景気判断理由文を説明変数，景気判断を被説明変数とした回帰モデルを学習し，景況感（野村 AI 景況感指数）の予測を行った．回帰モデルとしては，Long Short Term Memory（LSTM）ユニット [Hochreiter 97] から構成される両方向・多層の RNN（Recurrent Neural Network）を用い，学習・検証・テストデータとして，景気ウォッチャー調査の回答およそ 200,000 件を用いた．実験の結果，個別の景気判断の推定精度は平均二乗誤差（Mean Squared Error; MSE）で 0.346 と報告している．また，政府の月例経済報告と日銀の金融経済月報から抽出したテキストを同モデルを用いてそれぞれ指数化し，これらの結果が日銀短観 DI および景気ウォッチャー DI（現状）と高い相関（0.782〜0.843）にあることを示した．

類似の研究として，饗場ら [饗場 18] は Twitter から景気に関するツイートを抽出し，同様のモデルで景況感（SNS×AI 景況感指数）を予測した．近藤ら [近藤 19] は，金融機関の社員が取引先企業との面談を通して作成したテキストを入力として，同様のモデルで景況感を予測した．五島ら [五島 19] は，畳み込みニューラルネットを利用し，ロイターのニュース記事から景況感を予測した．

3 景況感の推定と要因の分析

2 節で述べたように，景気ウォッチャー調査にはアンケート回答者が判断した景気（現状および先行き）とその判断理由文とが含まれる．いずれも人手で作成された質の高い貴重な資源であり，これを基に経済指数を推定する試みが行われている．本研究では，五島ら [五島 19] と同様に新聞記事に着目して景況感指数の推定を行う．一方，五島らが景気ウォッチャー調査から学習したモデルにそのままニュース記事を入力して景況感指数を推定したのに対して，本研究では，3.1 節で述べるように記事のフィルタリング，およびモデルのドメイン適応を行う．なお，将来的には景況感指数のナウキャスト（足元予測）を目指しているが，本稿では既存の月次の景気指標（景気ウォッチャー DI）との比較のため，推定した指数を月次に集約して利用する．

続いて 3.2 節では，景況感に影響を与える要因の時間的な分析を行う手法について述べる．一般的に，景況感は，金融政策や物価，為替，雇用，賃金，海外情勢など，様々な要因から形成されている．しかしながら，全ての要因が経済に等しく影響を与えているわけではない．経済の情勢判断を行う上では，経済を上向きあるいは下向きに動かす要因が何であるのかを知ることが重要である．本研究では，どのような要因（語句）が,「いつ」,「どの程度」景況感に影響を与えたのか，推定した指数を基に分析を行う．

3.1 S-APIR 指数

新聞記事を利用した景況感指数（S-APIR）の予測を行うにあたり，必要となる 3 つのコンポーネントについて述べる．1 つ目は指数の予測自体を行う回帰モデル，2 つ目は新聞記事から無関係な記事を取り除くフィルタリング，最後に学習データ（景気ウォッチャー調査の景気判断説明文）と指数の予測に用いるデータ（新

聞記事）の違いを吸収するドメイン適応である．以下それぞれについて説明する．

3.1.1 回帰モデル

テキストデータを用いた回帰モデルでは，従来，各単語を独立した一つの説明変数としてモデルを学習することが一般的に行われてきた．そこでは，文脈を無視して，テキストは単語の集合（Bag of Words; BoW）として表現される [Manning 08]．しかしながら，テキストの意味をより適切に表現するためには，文脈による単語の意味の違いや単語間の係り受け関係を把握する必要がある．

近年，このようなテキストの文脈を考慮したモデル化の方法として，RNN がよく用いられている．ただし，RNN は勾配消失問題から長期の依存関係をうまく学習できないという欠点があるため，RNN の隠れ層を構成するニューロンとして，入力信号の依存関係を扱う特別な機構を持つ LSTM 等のユニットを用いるのが一般的である．本研究でも，テキストの文脈をより適切に表現するため，LSTM と同様の機構を持つ Gated Recurrent Unit（GRU）[Cho 14] を用いて RNN を構成する．そして，景気ウォッチャー調査のデータを学習データとして利用し，景気判断理由文を説明変数，景気判断の結果を被説明変数とする．ただし，現状判断と先行き判断とでは理由文の特徴が異なると考えられるため，本稿では現状判断のデータのみを用いる．また，景気判断は◎，○，□，△，× の記号で表現されているため，これらをそれぞれ 2, 1, 0, −1, −2 として扱う．

なお，先行研究 [山本 16] と同様に入力テキストの両方向から指数の推定を行う bidirectional LSTM（あるいは GRU），および近年注目されている Bidirectional Encoder Representations from Transformers（BERT）[Devlin 19] を利用した実験も行ったが，新聞記事から景気指標を推定するという本タスクにおいては，単純な GRU-RNN の方が総合的に良い結果が得られたため，本稿では他の結果は割愛する．

3.1.2 フィルタリング

前述のように学習した回帰モデルを利用することで，任意のテキストデータから景況感指数の予測が可能となる．しかしながら，本実験で利用するデータは新聞記事であるため，景気ウォッチャー調査の景気判断理由文とは異なり，入力される文章は経済に関係しないことも多い．そこで，経済や景気に関する記事だけを選択的に利用することを目的に，異常値検出に用いられる 1 クラス SVM（One-Class Support Vector Machine）[Manevitz 02] を利用する．学習データとして景気ウォッチャー調査の景気判断理由文を与えることで，景気判断理由文と類似していない文（新聞記事）を外れ値として指数の算出から除外する．

景気判断理由文の表現方法としては，tf-idf（term frequency-inverted document frequency）[Manning 08] で重み付けした単語ベクトル，BERT を用いた文の分散表現を比較し，予備実験でより良い結果が得られた前者を用いる．

3.1.3 ドメイン適応

RNN の学習に用いる景気ウォッチャー調査の景気判断理由文と，指数の推定に用いる新聞記事とでは，文章中の語彙や表現，文体（合わせて「ドメイン」と呼ぶ）が異なると考えられる．学習時と予測時におけるドメインの違いは，回帰モデルの予測性能に悪影響を与えるため，指数の推定に用いる新聞記事のドメインにモデルを適応させることができれば，より適切な指数の推定が可能になるものと期待できる．

そこで，景況感指数予測モデルのドメイン適応の可能性を探るため，処理対象の新聞記事から新たな学習データを自動生成し，ドメイン適応を試みる．具体的には，まず，景気ウォッチャー調査から学習した初期的な回帰モデル（3.1.1 節参照）を用いて新聞記事の各文の指数を予測する．そして，予測値が高い事例は景気動向をよく表していると仮定し，あらかじめ設定した閾値に基づいて，高スコアの文を正例，低スコアの（負で絶対値が大きい）文を負例として学習データとする．

このように構築した学習データを用いて，モデルのドメイン適応を行う．ドメイン適応は，ここでは単純に学習済みのモデルを新たな訓練データを用いて再学習（fine tuning）することで行う．

3.2 景況感に影響を与える要因の分析

景況感は，日銀の金融政策や消費増税などの税制，インバウンドや貿易・紛争などの海外情勢等々，様々な要因から形成されている．本節では，前節の方法で推定した指数に対する語の寄与を考えることで，どのような要因（語句）が，いつ，どの程度景況感に影響を与えたのかを分析する簡便な方法を提案する．具体的には，ある期間 t における語 w の寄与 $p_{t,w}$ を以下のように定義する．

まず，文 s の指数 p_s は，文 s を構成する語 $w \in s$ の指数 $p_{s,w}$ の総和であると仮定する．

$$p_s = \sum_{w \in s} tf_{s,w} \cdot p_{s,w} \tag{1}$$

$tf_{s,w}$ は，文 s における語 w の出現頻度である．さらに，すべての単語 $w \in s$ が等しく文 s の指数に貢献するというナイーブな仮定を導入すると，指数 $p_{s,w}$ は次式で表せる．

$$p_{s,w} = \frac{p_s}{|s|} \tag{2}$$

$|s|$ は文 s を構成する単語数である．ここで，期間 t 内の文の集合を S_t とする．そして，単語 w についてすべての文 $s \in S_t$ で指数 $p_{s,w}$ の総和を取り，総文数 $|S_t|$ で除することで，期間 t における語 w（のみ）の指数 $p_{t,w}$ を次式で定義する．

$$p_{t,w} = \frac{1}{|S_t|} \sum_{s \in S_t} tf_{s,w} \cdot \frac{p_s}{|s|} \tag{3}$$

直感的には，期間 t の記事に現れたすべての語 w の指数 $p_{t,w}$ の総和が，期間 t における S-APIR 指数であるというように解釈できる．

4 評価実験

4.1 データと実験設定

RNN モデルの学習のため，内閣府のホームページ[3]から，2000 年〜2018 年 9 月分の景気ウォッチャー調査の景気判断理由文（現状）計 216,741 件を収集した．そして，無作為に選んだ 90%を学習データ（検証データを含む），10%をテストデータとして用いた．なお，景気判断理由文は形態素解析器 MeCab[4]で分かち書きし，GRU-RNN への入力とした．GRU-RNN のパラメタは，学習データを用いた予備実験に基づき，GRU のユニット数を 512，RNN のレイヤー数を 2，語彙数を 40,000 とした．また，入力の単語は Wikipedia の記事であらかじめ学習された 300 次元の分散表現 [Bojanowski 17] を用いて表現した．

S-APIR 指数の予測には，日経新聞 2013〜2018 年度版（本紙朝刊および本紙夕刊）の記事見出しと記事本文を用いた．また，モデルのドメイン適応には，日経新聞 2010 年度版を用いた．いずれのデータも，各記事は句点「。」に基づいて文分割し，（記事単位ではなく）文単位で指数の予測を行った．モデル学習時と同様に，各文は MeCab を用いて形態素に分割してモデルへの入力とした．

4.2 景気ウォッチャー調査によるモデルの検証

景気判断理由文で学習した RNN を用いてテストデータで指数の予測を行い，実際の景気判断（−2 ∼ 2）との平均二乗誤差（MSE）を算出した．結果を表 2 に示す．比較のため，テキストデータを tf-idf で重み付けした単語ベクトルで表現し，リッジ回帰で指数の予測を行った結果も示す．リッジ回帰と比較し，RNN では誤差が大きく減少し，指数の予測がより正確に行えていることが確認できる．なお，この結果は山本ら [山本 16] の報告等とほぼ同様である．

表 2: 景気ウォッチャー指数の推定結果.

モデル	平均二乗誤差
tf-idf & リッジ回帰	0.509
GRU-RNN	0.351

4.3 景況感指数 S-APIR の予測

本節では，提案する景況感指数 S-APIR を既存の景気ウォッチャー DI と比較する．ただし，S-APIR 指数は，必ずしも景気ウォッチャー DI を置き換えることを目的としているわけではなく，ニュース記事をソースとした新しい景況感指数という位置づけである点に注意を要する．景況感には正解は無く，景気ウォッチャー DI も 2,050 名の限られた回答者の景気判断に基づく指数でしかない．本節での景気ウォッチャー DI との比較の目的は，あくまで S-APIR 指数がある程度既存の景況感指数と類似したトレンドを持っていることを確認すること，またどのような局面で既存の指数とは異なる動きをするのか，その特性を探ることにある．

4.3.1 実験結果

4.2 節で学習した RNN モデルを用いて，日経新聞 2013〜2018 年度版の記事見出しと記事本文を対象に景況感指数 S-APIR を算出した．なお，前述のように指数は文単位で推定されるため，発行月ごとにまとめて指数の推定値の平均を算出して景況感指数とした．比較のため，図 1 に，日経新聞から算出した景況感指数 S-APIR と内閣府発表の景気ウォッチャー DI（現状）を示す．おおまかには，S-APIR 指数は景気ウォッチャー DI のトレンドに近い動きを示しており，両者の相関係数も 0.546 と正の相関で見られた．

つづいて，1 クラス SVM で異常値検出を行い，景気ウォッチャー調査の景気判断理由に似ていない文を除外した上で，再度指数の推定を行った．結果を図 2 に示す．未だ景気ウォッチャー指数と異なる部分はあるものの，全体的にはさらに類似の動向を示すようになり，両者の相関係数は 0.546 から 0.686 に上昇した．この結果から，1 クラス SVM を用いたフィルタリングの効果が確認できる．

[3] http://www5.cao.go.jp/keizai3/watcher/watcher_menu.html
[4] http://taku910.github.io/mecab/

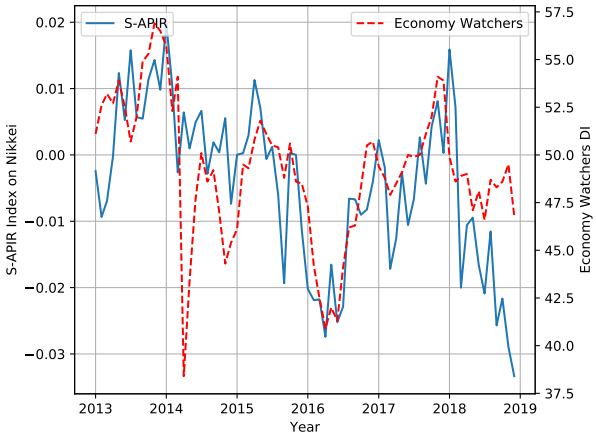


図 1: 日経新聞から算出した S-APIR 指数と景気ウォッチャー DI（現状）の比較（相関係数=0.546）.

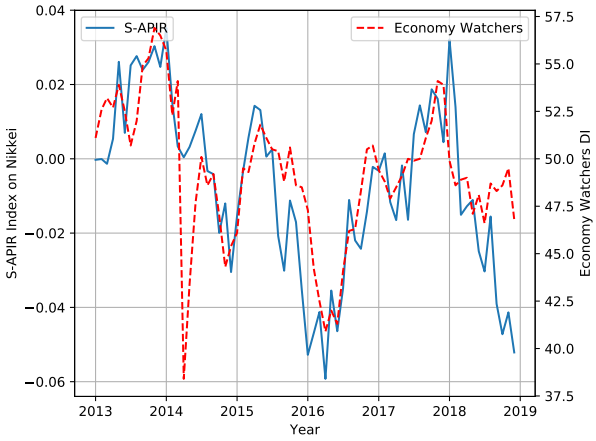


図 2: フィルタリング後の日経新聞から算出した S-APIR 指数と景気ウォッチャー DI（現状）の比較（相関係数=0.686）.

上で示したように，フィルタリングを用いて景気判断理由文と類似の文だけを対象に指数を算出することで，景気ウォッチャー DI との相関が顕著に向上した．しかしながら，景気判断理由文と新聞記事では文章の特徴が異なると考えられるため，前者で学習したモデルは後者には必ずしも適しているとは言えない．そこで，3.1.3 節で述べた方法でドメイン適応を試みた．具体的には，次の手順にしたがって実験を行った．

1. 日経新聞 2010 年版（本紙朝刊・本紙夕刊）の見出し・記事本文を抽出し，文ごとに MeCab で形態素に分割.
2. 1 クラス SVM を用いて異常値の文を検出・除去.
3. 学習済みの GRU-RNN で，残った各文の指数を予測.

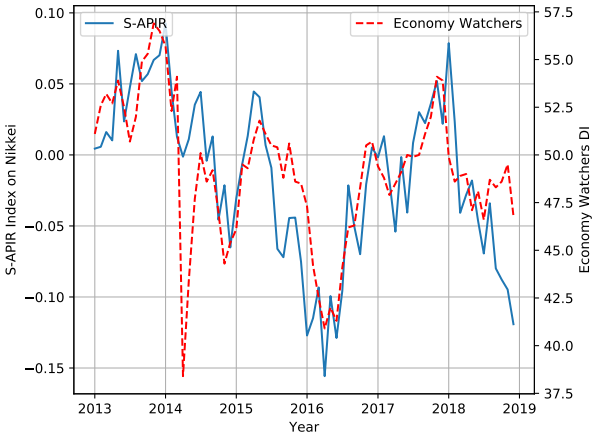


図 3: 回帰モデルの推定値によって構築した学習データでドメイン適応したときの景況感指数の推定結果（相関係数=0.701）.

4. 指数の絶対値が大きいものだけを学習データとして抽出．ここでは，指数の分布を目視で確認し，試みとして −1.0 以下に教師信号 −2，0.8 以上に教師信号 2 を与えて学習データとした.

この結果，正例（教師信号 =2）18,947 文，負例（教師信号 = −2）10,868 文のデータが学習データとして獲得できた．続いて，このデータを用いて，それぞれの新聞ごとにモデルの再学習を行った．そして，再学習したモデルを用いて再度指数の推定を行った．結果を図 3 に示す．ドメイン適応の結果，0.686 から 0.701 に相関係数が微増し，一定の効果が見られた．

なお，2014 年の前半に景気ウォッチャー DI は大きく落ち込んでおり，S-APIR 指数と大きな乖離がある．これは消費税が 8%に増税された時期であり，景気ウォッチャー DI と比較して，S-APIR 指数では増税の影響が小さいことは興味深い．この原因として，景気ウォッチャー調査の対象の約 7 割が家計動向関連業従事者であることが考えられる．すなわち，増税のように家計に与える影響が大きい要因はより強く景況感に影響を与えると考えられる．そこで，企業動向関連業従事者のみから算出された景気ウォッチャー DI（現状）と S-APIR 指数を比較したところ，相関係数は 0.819 と大きく上昇した．この結果は，日経新聞から算出した S-APIR 指数は，より企業動向を反映している指標であることを示している．

4.3.2 ドメイン適応によるモデルの変化

前節で，モデルの再学習を通したドメイン適応によって，S-APIR 指数と景気ウォッチャー DI の相関係数が向上することを実験的に示した．ここでは，モデル自

体が再学習によってどのように変化したのか調べるため，再学習前のモデルと再学習後のモデルそれぞれに対して，(新聞記事の文ではなく）単語のみをモデルの入力として与え，指数の予測を行った．表 3 に，それぞれのモデルで指数が大きかった 10 単語を上位から順に示す．上向きの矢印「↑」が付してある語は順位が上昇したもの，下向きの矢印「↓」は順位が下降したものを示す．「好調」,「享受」,「復調」などの語の順位が上昇する一方,「最高」,「絶好調」などの順位は降下している．

表 3: 単語のみで予測した景況感指数（上位 10 件）．

再学習前		再学習後	
快調	0.738	好調（↑）	1.591
絶好調	0.688	享受（↑）	1.503
好調	0.671	復調（↑）	1.502
最高	0.659	好転（↑）	1.488
上々	0.654	向上（↑）	1.485
好転	0.654	最高（↓）	1.484
伸	0.632	刺激（↑）	1.469
売れれ	0.630	絶好調（↓）	1.456
上向い	0.624	伸（↓）	1.449
着実	0.617	図り（↑）	1.437

同様に，表 4 で指数が小さかった 10 単語を比較する．「悪化」,「激減」などの順位が上昇する一方,「舗装」,「壊滅」の順位が下降している．モデルの再学習の結果，両者とも，ニュースで使われるような語彙が表の中でより上位に，アンケート回答で使われるような語彙がより下位にシフトしていることが確認できる．

表 4: 単語のみで予測した景況感指数（下位 10 件）．

再学習前		再学習後	
舗装	−1.543	悪化（↑）	−2.029
不通	−1.441	激減（↑）	−2.028
最悪	−1.355	最悪（−）	−2.026
壊滅	−1.334	舗装（↓）	−2.026
急変	−1.303	急変（−）	−2.019
全滅	−1.285	壊滅（↓）	−2.018
追い打ち	−1.282	低落（↑）	−2.017
激減	−1.278	減益（↑）	−2.017
エスカレート	−1.257	損失（↑）	−2.013
漁船	−1.232	急落（↑）	−2.011

なお，ここではモデルの再学習の効果を調べるために単語ごとの指数を予測したが，このようにして得られた語と指数の組は，単語とその極性値（感情値）とも見なせる．すなわち，この結果は経済ドメインの極性辞書としても利用することができる．一般ドメインにおいては Takamura ら [Takamura 05] の極性辞書など利用可能な資源がいくつか存在するものの，経済ドメインかつ日本語で自由に利用できる極性辞書は少ない．よって，この結果は，経済分野のテキストの感情分析に用いる言語資源としても有用であると考えられる．

4.4 景況感に影響を与える要因の分析

続いて，3.2 節に述べた方法で景況感に影響を与える要因について分析した．この方法では任意の語句について分析が可能であるが，ここでは例として「中国」,「貿易」について景況感指数への寄与を算出した結果を図 4，図 5 に示す．いずれも上側のグラフはドメイン適応したモデルで算出した S-APIR 指数であり，比較のために示している．

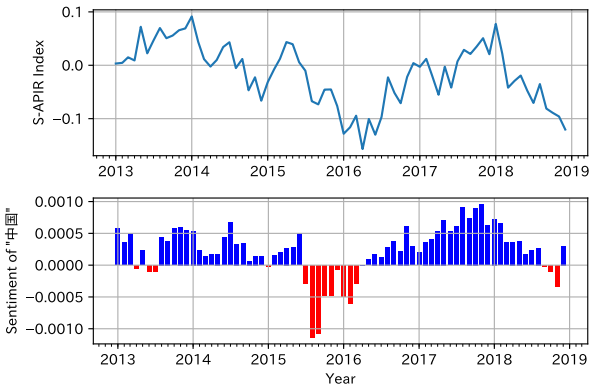


図 4: 景況感への「中国」の寄与．

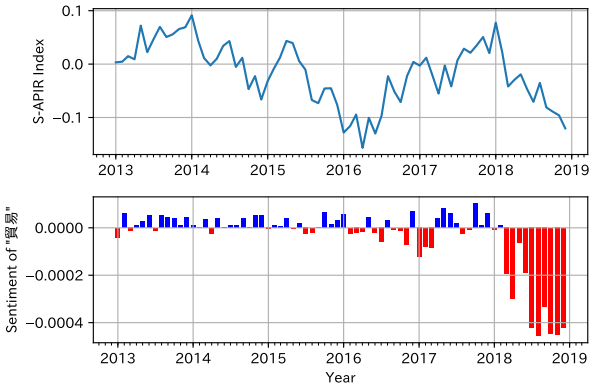


図 5: 景況感への「貿易」の寄与．

図 4 を見ると，S-APIR 指数と中国の寄与は似た動きをしており，中国に関する情勢が景況感に影響を与えていることが観察できる．特に，2015 年半ばから 2016 年初頭にかけては，中国の寄与度がマイナス側に大きく出ており，景況感を押し下げる一因となっていることが分かる．なお，この時期は，中国の景気が急速に悪化した時期である．

続いて図 5 を見ると，貿易に関する記事は，2013 年から 2018 年始めまで長らく景況感にほとんど影響を与えていなかったことが分かる．しかしながら，その後急速に寄与が低下し（マイナス方向に増加し），景況

感にマイナスの影響を与えている．これは，2018 年後半に始まった米中貿易摩擦を反映しているものと考えられる．

5　おわりに

本研究では，低コストで速報性の高い景況感指数を開発すること，および景況感に影響を与える要因を分析することを目的として，日々発信されるニュース記事を用いた指数予測の枠組みについて検討した．具体的には，1 クラス SVM による景気関連文の同定，GRU-RNN による指数の推定（回帰），新聞記事を用いたモデルのドメイン適応について実験的にそれぞれの効果を検証した．その結果，指数の予測に用いる入力を景気関連文に限定し，さらに新聞記事を用いてモデルを再学習することで，既存の景気ウォッチャー DI とより高い相関が得られることを確認した．なお，比較対象を企業動向に限定した場合，相関係数が大きく向上した．この結果は，本研究で開発した S-APIR 指数が，企業動向をより強く反映した指数であることを示唆している．そして，文単位で予測した指数を単語単位に分割して再集計することで，任意の要因（語句）が景況感に与える影響を時間軸に沿って分析できることをいくつかの例とともに示した．

謝辞

本研究は，一般財団法人アジア太平洋研究所（APIR）の自主研究プロジェクト「テキストデータを利用した新しい景況感指標の開発と応用」，JSPS 科研費 18K11558，平成 30 年度私立大学等経常費補助金特別補助「大学間連携等による共同研究」により行った．実験に使用した日経新聞データは APIR から提供を受けた．APIR の宮原秀夫氏，岩野宏氏，本多佑三氏，稲田義久氏，木下祐輔氏の御支援に感謝する．

参考文献

[饗場 18] 饗場 行洋, 山本 裕樹 : データサイエンスと新しい金融工学, 財界観測, Vol. 81, No. 2, pp. 30–41 (2018)

[Bojanowski 17] Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T.: Enriching Word Vectors with Subword Information, *Transactions of the Association for Computational Linguistics*, Vol. 5, pp. 135–146 (2017)

[Chakraborty 16] Chakraborty, S., Venkataraman, A., Jagabathula, S., and Subramanian, L.: Predicting Socio-Economic Indicators Using News Events, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1455–1464 (2016)

[Cho 14] Cho, K., Merrienboer, van B., Bahdanau, D., and Bengio, Y.: On the properties of neural machine translation: Encoder–decoder approaches, in *Proceedings of the 8th Workshop on Syntax, Semantics and Structure in Statistical Translation*, pp. 103–111 (2014)

[Devlin 19] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K.: 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*, pp. 4171–4186, Minneapolis, Minnesota (2019), Association for Computational Linguistics

[五島 19] 五島 圭一, 高橋 大志, 山田 哲也 : 自然言語処理による景況感ニュース指数の構築とボラティリティ予測への応用, 金融研究, Vol. 38, No. 3 (2019)

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, Vol. 9, No. 8, pp. 1735–1780 (1997)

[Kapetanios 18] Kapetanios, G. and Papailias, F.: Big Data & Macroeconomic Nowcasting: Methodological Review, Economic Statistics Centre of Excellence (ESCoE) Discussion Papers ESCoE DP-2018-12, Economic Statistics Centre of Excellence (ESCoE) (2018)

[近藤 19] 近藤 浩史, 與五澤 守, 成瀬 道紀, 森 正和 : 金融機関のテキストデータを活用した景気センチメントの計測, 人工知能学会全国大会論文集, pp. 1P2J1302–1P2J1302 (2019)

[Manevitz 02] Manevitz, L. M. and Yousef, M.: One-class SVMs for Document Classification, *The Journal of Machine Learning Research*, Vol. 2, pp. 139–154 (2002)

[Manning 08] Manning, C. D., Raghavan, P., and Schütze, H.: *Introduction to information retrieval*, Cambridge University Press, New York, NY, USA (2008)

[関 17] 関 和広, 水田 孝信, 八木 勲, 落合 友四郎, 酒井 浩之, 和泉 潔 : 金融情報学:ファイナンスにおける人工知能応用, 人工知能, Vol. 32, No. 6, pp. 905–910 (2017)

[Shapiro 17] Shapiro, , Hale, A., Sudhof, M., and Wilson, D.: Measuring News Sentiment, Federal Reserve Bank of San Francisco Working Paper 2017-01, Federal Reserve Bank of San Francisco (2017)

[Takamura 05] Takamura, H., Inui, T., and Okumura, M.: Extracting semantic orientations of words using spin model, in *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL2005)*, pp. 133–140 (2005)

[Watanabe 14] Watanabe, K. and Watanabe, T.: Estimating Daily Inflation Using Scanner Data: A Progress Report, in *CARF Working Paper Series*, No. CARF-F-342 (2014)

[山本 16] 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, 第 30 回人工知能学会全国大会 (JSAI 2016) (2016), 4 pages

景気の先行き判断の「先行き」の解釈における差異の識別: PU学習とテキストマイニングの応用

Identifying Different Definitions of Future in the Assessment of Future Economic Conditions: Application of PU Learning and Text Mining

加藤 真大 [1]

[1] 東京大学
[1] The University of Tokyo

Abstract: The *Economy Watcher Survey*, which is a market survey published by the Japanese government, contains *assessments of current and future economic conditions* by people from various fields. Although this survey provides insights regarding economic policy for policymakers, a clear definition of the word "*future*" in future economic conditions is not provided. Hence, the assessments respondents provide in the survey are simply based on their interpretations of the meaning of "future." This motivated us to reveal the *different interpretations of the future* in their judgments of future economic conditions by applying *weakly supervised learning* and *text mining*. In our research, we separate the assessments of future economic conditions into *economic conditions of the near and distant future* using *learning from positive and unlabeled data* (*PU learning*). Because the dataset includes data from several periods, we devised new architecture to enable neural networks to conduct PU learning based on the idea of *multi-task learning* to efficiently learn a classifier. Our empirical analysis confirmed that the proposed method could separate the future economic conditions, and we interpreted the classification results to obtain intuitions for policymaking.

1 Introduction

The *Economy Watcher Survey* is a market survey published by the Japanese government. The data consists of two types of assessments of economic conditions, *current and future economic conditions*, with five ranks. Although this survey provides policymakers with deep insights, it is difficult to interpret the assessments of future economic conditions because the meaning of *future* is not clearly defined and the definition thereof relies on the respondent's interpretation. Therefore, to obtain a clear understanding of survey participants' expectations, our approach was to classify assessments of future economic conditions into those pertaining to the *near and distant future*, respectively. This led us to propose a novel method that uses text data and a machine-learning algorithm in an attempt to grasp these expectations with respect to future economic conditions using data from the *Economy Watcher Survey*. For the classification task, we apply an algorithm that *learns from positive and unlabeled data* (*PU learning*), which is a machine-learning algorithm that enables us to train a classifier only from positive and unlabeled data.

Among studies of economic trends, methods using information contained in text data have become popular. Pioneering methods in this field are Tetlock [2007]; Tetlock *et al.* [2008], which involved the construction of sentimental indexes from articles of a column in the *Wall Street Journal* and an analysis of the predictability of the stock market. Kulkarni *et al.* [2009] predicted the residential price by using the number of searches on Google. Guzman [2011] also constructed real-time inflation expectations from search queries on Google.

PU learning is an algorithm of weakly supervised learning [Elkan and Noto, 2008; Ward *et al.*, 2009; Blanchard *et al.*, 2010; Nguyen *et al.*, 2011]. In the section describing the problem setting, we consider a situation in which only positive and unlabeled data

exist, and use only these data to train a binary classifier. PU learning has two scenarios known as *censoring scenario* and *case-control scenario* Elkan and Noto [2008]. In the study presented in this paper, we only focus on the *case-control* scenario, in which positive data are obtained separately from unlabeled data, and unlabeled data are sampled from the entire population. In this study, we construct our algorithm on the basis of subsequent research known as *unbiased PU learning* [du Plessis *et al.*, 2015], which minimizes the *unbiased estimator* of the classification risk.

After classifying the assessments of future economic conditions into those relating to the near and distant future, we calculated the averaged ranks for both the near and distant future. As a result, we found that a significant difference exists between economic conditions relating to these two future periods. This result infers the possibility that people's definition of the future differs. This fact is important from the viewpoint of economics. In macroeconomics, a researcher may be interested in the possibility of controlling people's expectations of the market. Our empirical analysis reports the fact that assessments of the economic conditions of the distant future were mainly based on economic fundamentals such as the population and diplomatic relationships.

In the following sections, we describe our problem setting and propose an algorithm that solves the problem. Subsequently, we present the results and interpretations of our empirical analysis.

2 Problem Setting

We consider the binary classification of text data. In the following parts, we describe the dataset and classification problem in detail.

2.1 Economy Watchers Survey

In our analysis, we used the *Economy Watchers Survey*, a dataset that contains text data and is published by the Japanese government [1)]. The purpose of this survey is to enable the region-by-region economic trends to be grasped accurately. This survey consists of two assessments, an assessment of current and future economic conditions with the possibility of entering sentences to motivate the answers by providing reasons. Respondents evaluated the current and future economic conditions by five ranks, $0, 1, 2, 3, 4$. The evaluation 0 means "worse" or "will get worse" compared with a previous period. The evaluation 4 means "better' or "will get better" compared with a previous period. The evaluation 2 represents a neutral position on the assessment of economic conditions.

Interpretation of Assessment of Future Economic Conditions: Assessments of current and future economic conditions provide us with deep insights into economic reality. However, in the questionnaire, there is no clear definition of the concept of the "future" with respect to future economic conditions. Hence, different people interpret the duration of "future" in their own way. Whereas one person may imagine the future as just one week, the "future" might be a few months for another person. Therefore, to analyze the assessments more accurately, we need to classify assessments of future economic conditions as being either near or distant economic conditions.

2.2 Classification of an Assessment of Future Economic Conditions

To classify future economic conditions into those expected to occur in either the near or distant future, we propose assuming that current economic conditions share similar sentences with those expected in the near future. Our classification strategy is to regard current economic conditions as positive data and future economic conditions as unlabeled data, which potentially consists of positive and negative data. In this paper, positive data are assessments of the current economic conditions and those expected in the near future, whereas negative data are assessments of economic conditions foreseen to prevail in the distant future. We illustrate the relationship between assessments of current and future economic conditions of our assumption on Figure 1. We train our classifier only from positive and unlabeled data by using an algorithm that employs PU learning. Therefore, the goal of this problem is to classify $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$ into one of the two classes $\{-1, +1\}$, where $+1$ denotes assessments of current economic conditions and those expected in the near future (positive data) and -1

[1)] Particulars of the dataset are provided on the homepage of the Japanese government, `https://www5.cao.go.jp/keizai3/watcher-e/index-e.html`. The survey enlists the cooperation of people holding employment positions that enable them to observe activity closely related to the regional economy. We downloaded the dataset from the page.

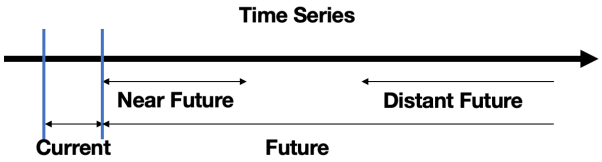


図 1: Our assumed definition of the time structure of assessments.

denotes economic conditions relating to the distant future (negative data).

2.3 Data Generating Process of Economy Watchers Survey

Let us describe the data generating process of our problem. Let us assume that we have n data points at t-th period and denote the i-th text data as $\boldsymbol{x}_i \in \mathcal{X} \subset \mathbb{R}^d$. If the target of text data $\boldsymbol{x}_i$ describes current or near future economic conditions, we attach a positive label, i.e., $y_i = +1$. If the target of text data $\boldsymbol{x}_i$ describes distant future economic conditions, we attach a negative label, i.e., $y_i = -1$. However, in the dataset, we can only observe positive data, and unlabeled data, which includes both positive and negative data. In addition, if the text data $\boldsymbol{x}_i$ belongs to a period $t \in \{1, ..., T\}$, we denote the fact as $z_i = t$. Using these notations, we define our data generating process as follows:

$$\{\boldsymbol{x}_i\}_{i=1}^{n} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x}|y=+1, z=t),\ \{\boldsymbol{x}'_i\}_{i=1}^{n'} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x}|z=t),$$

where $\{\boldsymbol{x}_i\}_{i=1}^{n}$ and $\{\boldsymbol{x}'_i\}_{i=1}^{n'}$ denote the positive and negative data at t-th period, and $p(\boldsymbol{x}|z=t)$ can be decomposed as

$$\begin{aligned} p(\boldsymbol{x}|z=t) =& p(y=+1|z=t)p(\boldsymbol{x}|y=+1, z=t) \\ &+ p(y=-1|z=t)p(\boldsymbol{x}|y=-1, z=t). \end{aligned}$$

3 Learning from Positive and Unlabeled Data with Time Series Data

To classify data consisting only of positive and unlabeled data, we propose using *multi-task PU learning* (*MTPU*). In this section, we provide details of the proposed algorithm.

3.1 Learning from Positive and Unlabeled Data

Before explaining our model, let us explain the standard setting of PU learning. In PU learning, we consider a binary classification problem to classify $\boldsymbol{x} \in \mathcal{X} \subset \mathbb{R}^d$ into one of the two classes $\{-1, +1\}$. We assume that there exists a joint distribution $p(\boldsymbol{x}, y)$, where $y \in \{-1, +1\}$ is the class label of $\boldsymbol{x}$. PU learning relies on two distinct sampling schemes, namely the censoring scenario and case-control scenario [Elkan and Noto, 2008]. The PU learning framework we use in this study is the case-control scenario, in which we suppose access to a positive dataset $\{\boldsymbol{x}_i\}_{i=1}^{n} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x}|y=+1)$ and an unlabeled dataset $\{\boldsymbol{x}'_i\}_{i=1}^{n'} \overset{\text{i.i.d.}}{\sim} p(\boldsymbol{x})$. Let $\ell : \mathbb{R} \times \{\pm 1\} \to \mathbb{R}^+$ be a loss function, where $\mathbb{R}^+$ is the set of non-negative real values, and $\mathcal{F}$ be the set of measurable functions from $\mathcal{X}$ to $[\epsilon, 1-\epsilon]$, where $\epsilon \in (0, 1/2)$ is a small positive value. This constant ϵ is introduced to ensure the following optimization problem is well-defined based on the result of Kato *et al.* [2019]. Here, du Plessis *et al.* [2015] showed that the classification risk of $f \in \mathcal{F}$ can be expressed as

$$\begin{aligned} R_{\mathrm{PU}}(f) =& p(y=+1)\mathbb{E}_{\mathrm{p}}[\ell(f(X), +1)] \\ &- p(y=-1)\mathbb{E}_{\mathrm{p}}[\ell(f(X), -1)] + \mathbb{E}_{\mathrm{u}}[\ell(f(X), -1)], \end{aligned} \tag{1}$$

where $\mathbb{E}_{\mathrm{p}}$ and $\mathbb{E}_{\mathrm{u}}$ are the expectations over $p(\boldsymbol{x}|y=+1)$ and $p(\boldsymbol{x})$, respectively. The above formulation of PU learning provides the unbiased risk of the classification problem.

3.2 Multi-Task Non-negative PU learning for Time Series Data

In addition to the standard setting of PU learning, we could take the time structure into account. The Economy Watcher Survey comprises monthly data, with approximately $2,600$ records for each month. Here, we would need to use different classifiers for the data included in each month for the following two reasons. First, the model can vary across periods. Second, we would not be able to include data of the $(t+1)$-th period to train a model of data of the t-th period because the data of the $(t+1)$-th period might have information of the data of the t-th period. This made it necessary to use different models across different periods. For $z = t$, we denote the model as

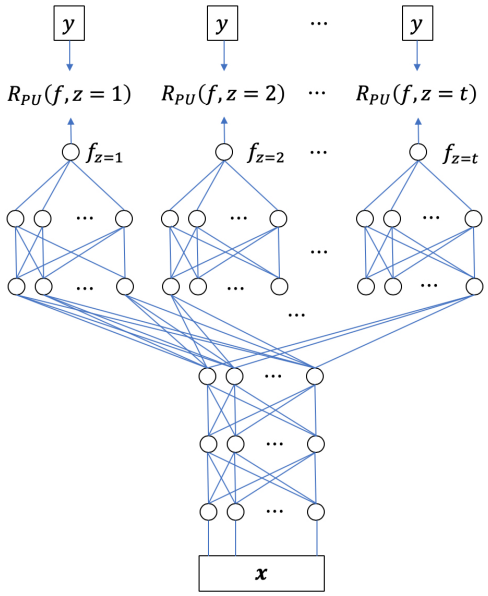


図 2: Neural network model for multi-task learning for PU learning. The models share one shared network with 3 layers.

$f_{z=t}$ and the risk as follows:

$$R_{\mathrm{PU}}(f_{z=t}, z=t) = p(y=+1|z=t)\mathbb{E}_{\mathrm{p,t}}[\ell(f(X),+1)] - p(y=-1|z=t)\mathbb{E}_{\mathrm{p,t}}[\ell(f(X),-1)] + \mathbb{E}_{\mathrm{u,t}}[\ell(f(X),-1)],,$$

where $\hat{\mathbb{E}}_{\mathrm{p,t}}$ denotes the averaging operator over positive data, $\hat{\mathbb{E}}_{\mathrm{u,t}}$ denotes averaging over the unlabeled data at the t-th period. We additionally introduce a model for *multi-task learning* to PU learning. Multi-task learning is proposed to train neural networks efficiently by using the common features across different tasks Caruana [1997]. If a common feature exists across periods, we can train our models more efficiently by sharing the common feature among models $f_{z=t}$ for $t = 1, ..., T$ through the layers named *shared layers*, the structure of which is shown in Figure 2. We named this model *MTPU*. Details of its structure are provided in the section for empirical experiments.

3.3 Sample Approximation of the Unbiased Risk

When we train a classifier, we can naively replace the expectations with the corresponding sample averages. However, Kiryo *et al.* [2017] pointed out that the basic form of the unbiased PU learning is ineffective with a deep neural network because of overfitting caused by the fact that the risk is not lower bounded. To implement PU learning with deep neural networks, we applied the non-negative risk proposed by Kiryo *et al.* [2017] to the empirical risk defined in (2). For a hypothesis set $\mathcal{H}$, let us define the following risk minimization problem,

$$\hat{f}_{z=t} = \mathrm{argmin}_{f_{z=t}\in\mathcal{H}} \left[\widehat{R}_{\mathrm{PU}}(f_{z=t}, z=t) + \mathcal{R}(f)\right], \tag{2}$$

where $\widehat{R}_{\mathrm{nnPU}}(f_{z=t}, z=t)$ is a sample approximation of $R_{\mathrm{PU}}(f_{z=t}, z=t)$ with non-negative transformation proposed by Kiryo *et al.* [2017] and $\mathcal{R}$ is a regularization term.

3.4 Class Prior and Selection Bias

The remaining problem is to make a decision regarding the class prior $p(y=+1|z=t)$. The class prior $p(y=+1|z=t)$ would be different across periods t. Although several algorithms have been proposed to estimate the class prior [du Plessis and Sugiyama, 2014; Ramaswamy *et al.*, 2016; Jain *et al.*, 2016], the estimation is still known to be a difficult task. However, we can avoid the problematic estimation in the case of the particular goal we hope to reach. In our experiments, we assume that the class prior is $p(y=+1|z=t) = 0.2$ for all periods, $t = 1, 2..., T$. This assumption is not realistic because the probability would have different values across the periods. However, Kato *et al.* [2018, 2019] showed that the function $f_{z=t}$ is simply linear-proportional to the class prior, i.e., the following relationship holds even if we miss-specify the class prior:

$$\begin{aligned} &p(y=+1|x, z=t) \le p(y=+1|x, z=t) \\ &\Leftrightarrow f_{z=t}(x) \le f_{z=t}(x). \end{aligned} \tag{3}$$

Therefore, even when we cannot obtain the exact value of $p(y=+1|\boldsymbol{x}, z=t)$, we can still identify the order of $p(y=+1|\boldsymbol{x}, z=t)$ with regard to $\boldsymbol{x}$. Our empirical analysis separates the assessment of future economic conditions into near and distant future economic conditions based on this property. We classify 1/5 of data from the highest value of $f_{z=t}$ into assessments of near future economic conditions, and 1/5 of data from the lowest value of $f_{z=t}$ into assessments of distant future economic conditions. In addition to the robustness to the miss-specified class prior, the function $f_{z=t}$ also holds the relationship 3 under the selection bias of positive data [Kato *et al.*, 2019] if our assumption is

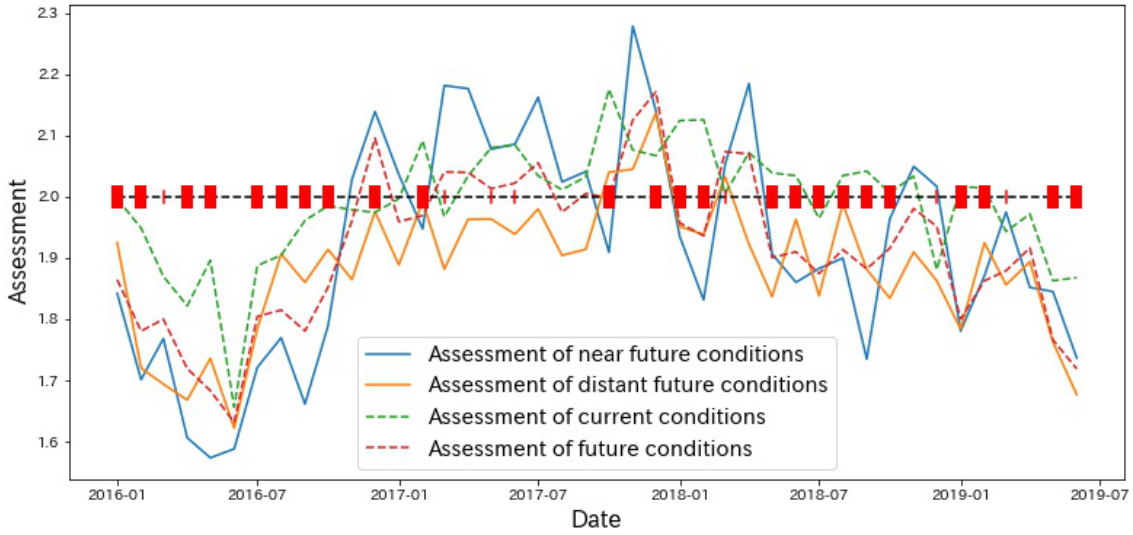


図 3: Plotted assessments of the economic conditions of the near and distant future and those pertaining to the present and future. The horizontal line at $y = 2$ is the neutral state. The red vertical lines on the horizontal line represent the results of the two-sample t-test. The thin and bold red vertical lines represent the 5% and 1% significance levels, respectively.

mild. Thus, our results can reduce the influence of the miss-specified class prior and selection bias.

4 Empirical Analysis

In this section, we report the results of the empirical analysis of data from the Economy Watcher Survey. The survey was conducted every month starting in 2000. Our analysis only used data from January 2016 to June 2019, i.e., 42 months' data. Each month includes approximately 2,600 samples. The reason for the heterogeneity among the data is the lack of text in respondents' answers. In total, we had 111,501 samples.

We used Bag-of-Words to represent the documents as 16,914-dimensional vectors. After vectorizing the text data, we applied PU learning with the aforementioned MTPU. In addition to the model, we also used the standard model of PU learning to compare the performance. We used this model of PU learning in two ways. First, we used all samples to train one model. Second, we prepared one model for each month. Details of the neural networks are provided in the following section. After training our classifier, we classified the assessment of future economic conditions using unlabeled data that we used for training.

Neural network model: First, we describe the model we used for MTPU. The model for the shared network was a 3-layer *multilayer perceptron* (MLP) with ReLU Nair and Hinton [2010] (more specifically, $16914 - 500 - 500 - 500$). The neural network model following the shared network was a 2-layer MLP (more specifically, $500 - 500 - 1$) with ReLU. Next, we describe the model we used for non-negative PU learning. The model for the neural network was a 5-layer MLP (more specifically, $16914 - 500 - 500 - 500 - 500 - 1$) with ReLU. We set $p(y = +1|z = t) = 0.2$ for all $t \in \{1, 2, ..., 42\}$. For both methods, we use logistic loss for the loss function ℓ.

4.1 Difference among the Assessments

In this section, we report the extent to which assessments differ across current, future, near future, and distant future economic conditions.

Averaged Assessments and t-test: We report the averaged assessments of economic conditions in the

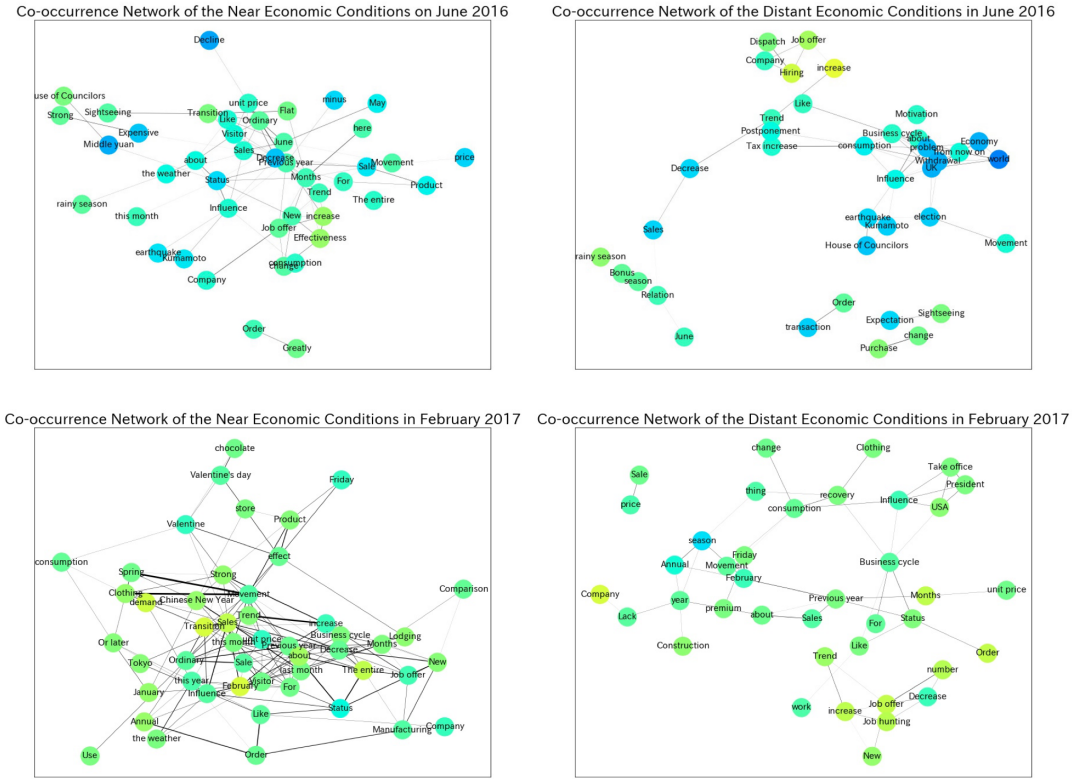


図 4: Co-occurrence network of near and distant future economic conditions in June 2016 and February 2017. The lengths of the edges represent the value of the Jaccard coefficients. Shorter edges indicate a stronger relationship (the value of the Jaccard coefficients is larger) between the two words. The widths of the edges also represent the value of the Jaccard coefficients between the two words. The bold edges similarly signify a stronger relationship (the value of the Jaccard coefficients is larger) between the two words. The color of the nodes relates to the assessment. The yellow-green color denotes that the averaged value is 2, i.e., the assessment is neutral. The warmer and cooler colors represent positive and negative assessments, respectively.

near and distant future in comparison with those of the current and future. Assessments of the near and distant future economic conditions are estimated by MTPU and non-negative PU learning with neural networks. For non-negative PU learning, we used two models. The first (named PU1) entailed training one model for all samples. The second (named PU2) involved using different models for the data of different months. The results are presented in Table 1. For each period, we show the results of the two-sample t-test with unequal variances between the assessments of the economic conditions of the near and distant future. Values for which the difference between the mean of the assessments is significant are indicated by superscript $*$ in the table. One $*$ and two $*$s mean that the null hypothesis of the two-sample t-test is rejected at the 5% and 1% significance levels, respectively.

Visualization as a Time Series: To facilitate a more intuitive understanding of the reported results, we plotted the averaged assessments in the time series in Figure 3, where the x-axis corresponds to the time series, and the y-axis corresponds to the value of the assessment. The blue, orange, green, and red lines correspond to assessments of the economic conditions in the near future, distant future, at the present time, and in the future. The horizontal black dashed line at $y = 2.0$ represents the neutral condition in the 5-step evaluations for the economic conditions from 0 (bad) to 4 (good). The vertical red lines perpendicular to the line $y = 2.0$ indicate that the difference between the average assessments of the economic conditions in the near and distant future is significant in the two-sample t-test. The bold vertical lines represent that the null hypothesis of the two-sample t-test is rejected at the 1% significance level and the other red lines

represent that two $*$ means that the null hypothesis of the two-sample t-test is rejected at the 5% significance level. For example, the assessments of the economic conditions in the near future in 2017 are significantly higher than those of the distant economic conditions.

4.2 Text Mining

This section presents our analysis of the text based on assessments of the text data. For text mining, we use *tf-idf* and the *Jaccard coefficient*, which are standard techniques of natural language processing. First, we separate the assessments of the economic conditions in the near and distant future for the month in which the assessments were published, i.e., we form groups of monthly assessments. Then, we denote a set of the group of assessments as $\mathcal{M}$, and we apply tf-idf to identify the word that characterizes the document. Then, for the 50 words with the highest tf-idf, we measure the Jaccard coefficient Manning and Schütze [1999], which measures the similarity between two sets. Let $\mathcal{M}_w \in \mathcal{M}$ be a set of sentences including the word w. The Jaccard coefficient $J(\mathcal{M}_a, \mathcal{M}_b)$ for a word a and a word b can be expressed as follows:

$$J(\mathcal{M}_a, \mathcal{M}_b) = \frac{|\mathcal{M}_a \cap \mathcal{M}_b|}{|\mathcal{M}_a \cup \mathcal{M}_b|}. \tag{4}$$

Based on these results, we plotted the co-occurrence networks in Figure 4[2)]. Because of the limitation placed on the length of the paper, we only show the network of assessments in June 2016 and February 2017. June 2016 is one of the periods in which the value of assessments greatly changed. Throughout 2017, the economic conditions of the near future are less than those in the distant future, and February 2016 is one of these periods. Because of the small size of our graphs, we placed enlarged versions of these graphs in the appendix[3)] in both English and Japanese.

4.3 Interpretations

Figure 4 displays words related to economic fundamentals, such as the structure of the labor supply and international politics. In other words, these results can be interpreted as meaning that assessments of the economic conditions of the near future represent the economic cycle, whereas assessments of the economic conditions of the distant future represent the economic trend. For example, the words "U.K." and "withdrawal" appear, both of which are related to Brexit among the economic conditions of the distant future, "Business cycle," and "Trend" in June 2016. The words "US" and "President" appear in Feb. 2017. On the other hand, the economic conditions of the near future in June 2016 and Feb, 2017 are represented by words that have less relationship with economic fundamentals such as "rainy season" and "Valentin's day." For policymakers, this is an insightful finding because the result infers that they cannot easily change people's expectations based on economic fundamentals.

5 Conclusion

In this paper, we proposed a new application of PU learning and text mining to data consisting of financial text. We developed a new model named MTPU to train neural networks efficiently using data with a time structure. Our empirical analysis showed the classification result and interpretations based on text mining and economics. The result is insightful to policymakers because the result infers that people might have a different interpretation of the definition of the future and may assess the future economic outlook differently based on their interpretations of the future. Besides, we also found that there are different main reasons between near and distant future economic assessments.

参考文献

Gilles Blanchard, Gyemin Lee, and Clayton Scott. Semi-supervised novelty detection. *Journal of Machine Learning Research*, 11(Nov):2973–3009, 2010.

Rich Caruana. Multitask learning. *Machine Learning*, 28(1):41–75, Jul 1997.

Marthinus Christoffel du Plessis and Masashi Sugiyama. Class prior estimation from positive and unlabeled data. *IEICE Transactions on Information and Systems*, E97-D(5):1358–1362, 2014.

[2)] We translated from Japanese to English using an API provided by Google (`https://pypi.org/project/googletrans/`).

[3)] https://arxiv.org/abs/1909.03348

Marthinus Christoffel du Plessis, Gang. Niu, and Masashi Sugiyama. Convex formulation for learning from positive and unlabeled data. In *ICML*, pages 1386–1394, 2015.

Charles Elkan and Keith Noto. Learning classifiers from only positive and unlabeled data. In *ICDM*, pages 213–220, 2008.

Giselle Guzman. Internet search behavior as an economic forecasting tool: The case of inflation expectations. *Journal of Economic and Social Measurement*, 36, 11 2011.

Shantanu Jain, Martha White, Michael W Trosset, and Predrag Radivojac. Nonparametric semi-supervised learning of class proportions. In *NIPS*, 2016.

Masahiro Kato, Liyuan Xu, Gang Niu, and Masashi Sugiyama. Alternate estimation of a classifier and the class-prior from positive and unlabeled data. *arXiv:1809.05710*, 2018.

Masahiro Kato, Takeshi Teshima, and Junya Honda. Learning from positive and unlabeled data with a selection bias. In *International Conference on Learning Representations*, 2019.

Ryuichi Kiryo, Gang Niu, Marthinus Christoffel du Plessis, and Masashi Sugiyama. Positive-unlabeled learning with non-negative risk estimator. In *NIPS*, pages 1675–1685, 2017.

Rajendra Kulkarni, Kingsley Haynes, Roger Stough, and Jean Paelinck. Forecasting housing prices with google econometrics. *SSRN Electronic Journal*, 07 2009.

Christopher D. Manning and Hinrich Schütze. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, MA, USA, 1999.

Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *ICML*, 2010.

Minh Nhut Nguyen, Xiaoli-Li Li, and See-Kiong Ng. Positive unlabeled leaning for time series classification. In *IJCAI*, pages 1421–1426, 2011.

Harish Ramaswamy, Clayton Scott, and Ambuj Tewari. Mixture proportion estimation via kernel embeddings of distributions. In *ICML*, pages 2052–2060, 2016.

Paul C. Tetlock, Maytal Saar- tsechansky, and Sofus Macskassy. More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance*, 63(3):1437–1467, 2008.

Paul C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168, 2007.

Gill Ward, Trevor Hastie, Simon Barry, Jane Elith, and John R Leathwick. Presence-only data and the em algorithm. *Biometrics*, 65(2):554–563, 2009.

アナリストレポートにおける キーワード関連文の抽出と景況感推移観測への応用

Extraction of Keyword Related Sentences in Analyst Reports and Its Application to Observation of Business Confidence

高山 将丈[1]　小澤 誠一[1,2*]　廣瀬 勇秀[3]
飯塚 正昭[3]　渡辺 一男[3]　逸見 龍太[3]
Shota TAKAYAMA[1]　Seiichi OZAWA[1,2]　Takehide HIROSE[3]
Masaaki IIZUKA[3]　Kazuo WATANABE[3]　Ryuta HENMI[3]

[1] 神戸大学大学院 工学研究科
[1] Graduate School of Engineering, Kobe University
[2] 神戸大学 数理・データサイエンスセンター
[2] Center for Mathematical and Data Sciences, Kobe University
[3] 三井住友 DS アセットマネジメント株式会社
[3] Sumitomo Mitsui DS Asset Management Company, Limited

Abstract: Analysts in investment trust management companies survey business achievements of target companies and are supposed to summarize them in the form of an analyst report. A fund manager reviews the reports and decides which companies to invest based on the reports as well as various economic indices and information. However, when searching for potential investment targets, a fund manager is generally required to read a large number of analyst reports and other related documents, Obviously, this work is not an easy task even for a skilled manager. In this work, we propose an intelligent system that retrieves meaningful sentences related to a specific query such as 'performance' and automatically evaluates a market trend in order to mitigate their work loads. From a total of 37,398 analyst reports and interview records, we obtained word embedding vectors using Word2Vec, and related sentences addressing company's financial soundness were retrieved based on the similarity to a query. In our experiments, for the word 'achievement', we retrieved 395 sentences out of 2,182 sentences that were 2.67 times larger than those when an exact search was applied. On the other hand, a trend of market sentiment obtained from a keyword such as 'profit' did not have high correlation against actual market indices.

1　はじめに

投資信託を請け負う投資信託運用会社（運用会社）は，投資によって利益が得られる優良企業を見出し，その投資額を適切に設定することが求められる．このため，運用会社に所属するアナリスト及びファンドマネージャーは，日常的に投資候補企業の調査を行い，投資先の選定を行っている．具体的には，まずアナリストが投資候補企業の調査を行い，その調査結果を速報的にアナリスト往訪記録という文書にまとめ，十分に情報収集ができたタイミングで有価証券報告書や株価，同業他社の情報なども分析して，正式な報告書であるアナリストレポートを発行する．このアナリストレポートをもとに，ファンドマネージャーは投資候補企業への投資判断を行っている．この投資先決定フローにおいて，ファンドマネージャーは，アナリストがまとめた多数のレポートを精読しなければならず，これには多大な労力と時間を要する．そこで，この業務をサポートするシステムの開発が求められている．

このようなシステムとして，深層学習を用いたデータ駆動型の文書解析手法が注目されており，アナリストレポートの文書表現から景況感を読み取る試みが行われている．小林ら [1] は，アナリストレポートからアナリスト予想の根拠を含む文を自動抽出できること

*連絡先：神戸大学 数理・データサイエンスセンター
〒 657-8501 神戸市灘区六甲台町 1-1
E-mail: ozawasei@kobe-u.ac.jp

を示した．また，得られた文に対して極性付与を行い，75.13%の精度で極性を推定できることを示した．小林らの手法では，「手がかり表現」と呼ばれるキー単語を人手で入力して，アナリスト予想の根拠となる文の特徴抽出が試みられている．そして，得られた特徴量に対して 12 層の全結合ニューラルネットワークを利用し，景況感判定を行っている．この手法では，根拠情報を含む文しか出力できないことや事前に定義した「手がかり表現」を入力する必要があり，分析者がアナリストレポートの特徴を熟知していることが前提となる．また，多数の全結合層で構成されたニューラルネットワークを使用しているため，勾配消失や学習の発散，過学習などが懸念される．

本研究では，分析者の専門知識を前提としなくても自動的に手がかり表現の関係性が学習され，指定したキー単語に依存しない関連文抽出が可能となる深層学習モデルのアプローチを採用する．ニューラルネットの構造を制約することで，全結合ニューラルネットワークで問題となる勾配消失等の問題を緩和し，文書の分類タスクで著しい成果を挙げている「BERT[2]」を用いたシステムを提案する．

2　提案システム

本論文では，アナリストレポートにおいて，蓄積されたアナリストレポートを用いて，入力された視点に対する景況感の推移を確認できるシステムを提案する．

本システムの概略図を図 1 に示す．本提案システムは大きく分けて 2 つのモジュールで構成されており，それぞれ図 1 の，課題 2 で示されている部分である．1 つ目は Word2Vec[3] を用いて入力したキーワードにおける関連文をアナリストレポートから抽出するモジュールである．2 つ目は自然言語処理の各種タスクにおいて目覚ましい成果を挙げている BERT を用いたモジュールで，1 つ目のモジュールで得られた関連文をもとにキーワードにおける景況感推移を観測するものである．

2.1　キーワード関連文抽出モジュール

例えば，「業績」という観点で景況感追跡を行いたい場合，「業績」という単語を持つ文を抽出するだけでは不十分であることは想像に難くない．「業績」に関して述べるために「決算」や「収益」など，「業績」に近い意味を持つ単語を用いている可能性も低くないからである．

本モジュールの概略図を図 2 に示す．これは「業績」だけでなく「決算」や「収益」などの単語を含んだ文の抽出を行うことを目的とした．本モジュールに入力した「業績」などのキーワードと近い意味を持つ「決算」や「収益」などの単語を，Word2Vec を用いて抽出し，それらの単語を含む文を関連文としてアナリストレポートから抽出することによって実現する．

まずアナリストレポート及びアナリスト往訪記録をコーパスとして学習させた Word2Vec を用いて，入力したキーワードと近い意味を持つ単語 (関連語) を抽出する．上述の例においては，キーワードが「業績」，関連語が「決算」や「収益」となる．そして，キーワードと関連語を含む文を抽出すること本モジュールの出力とする．

このようにして得られたキーワード関連文は，次の景況感推移観測モジュールに入力される．

2.2　景況感推移観測モジュール

本モジュールは，自然言語処理の各種タスクにおいて目覚ましい成果を挙げている BERT を用いたモジュールで，キーワード関連文抽出モジュールで得られた関連文をもとにキーワードにおける景況感推移を観測するものである．概略図を図 3 に示す．

まず Wikipedia コーパスを用いて BERT の事前学習を行う．その後景況感ラベルの付与されたアナリスト文書を用いて BERT の Fine-Tuning を行い，文章から景況感を推定できるように学習を行う．これによって得られた学習済み BERT モデルを用いて，キーワード関連文抽出モジュールから入力された文ごとに景況感判定を行い，年度内で景況感を平均する．これによってキーワードにおける年度ごとの景況感が抽出可能となる．

3　実験設定

3.1　検証に利用したデータ

本提案システムは特定企業におけるキーワード別の景況感を追跡するものである．本検証では東証一部上場企業 1 社について検証を行った．この企業を企業 AB と定義する．存在したアナリストレポートは，2002 年から 2016 年に発行されたもので，総文書数は 76 文書，総文数は 1764 文であった．アナリストレポートは docx ファイルで発行されていたため，Python モジュールである docx2txt[1] を利用し，文字列の抽出を行った．

3.2　キーワード関連文抽出モジュール

三井住友 DS アセットマネジメント株式会社から提供を受けたアナリストレポート及びアナリスト往訪記

[1] docx2txt · PyPI, https://pypi.org/project/docx2txt/

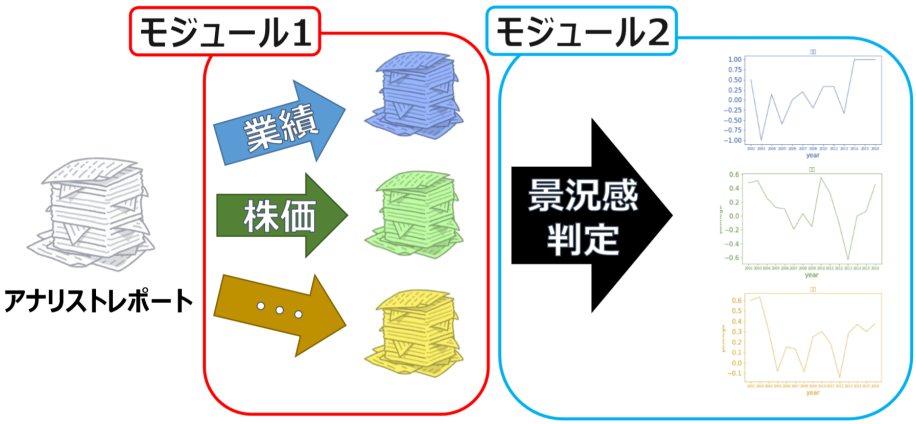


図 1: 提案手法の概略図

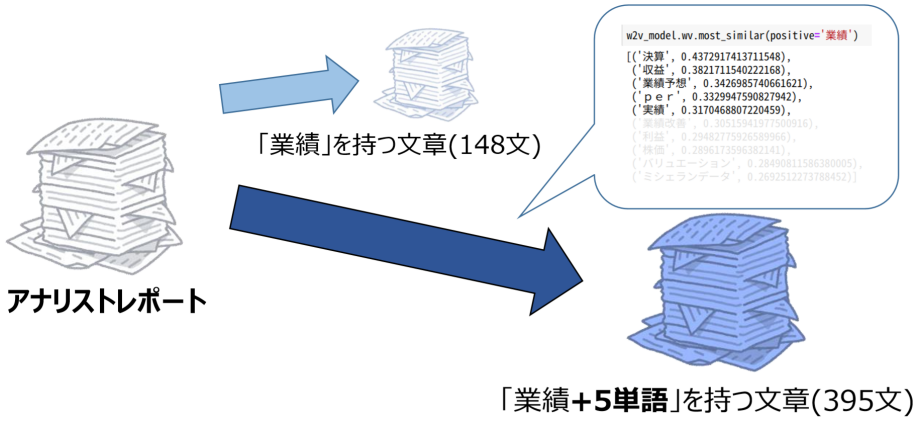


図 2: キーワード関連文抽出モジュールの概略図

表 1: アナリスト文書の内訳

文書の種類	ラベル	数量
アナリスト往訪記録	ポジティブ	1,145
	ネガティブ	788
	ニュートラル	2,160
	-	9,580
アナリストレポート	-	21,892

表 2: Word2Vec のパラメーター

パラメーター名	値
ウィンドウサイズ	2
単語ベクトルの次元数	500
min_count	1
イテレーション数	20

録をコーパスとして，Word2Vec の学習を行った．利用した文書データの内訳及び Word2Vec の学習のパラメーターをそれぞれ表 1 及び表 2 に示す．

また学習前の処理として，これらの文書に適切な前処理を行った後，活用する単語の原形への変換，数値データの正規化，記号の削除の処理を行った．分かち書きには Python モジュールである MeCab を利用し，その辞書として NEologd[4] を利用した．

学習後の Word2Vec を用いてキーワードに対する関連語抽出するが，本検証では Word2Vec の出力する単語ベクトルのうち，キーワードのベクトルにコサイン類似度の近い上位 5 つの単語を関連語と定義した．

3.3 景況感推移観測モジュール

本検証ではウェブ上で一般公開されている事前学習済み BERT モデルを利用した[2]．これは Wikipedia コーパスを用いて学習させたもので，SentencePiece[5] で

[2]BERT with SentencePiece を日本語 Wikipedia で学習してモデルを公開しました，https://yoheikikuta.github.io/bert-japanese/

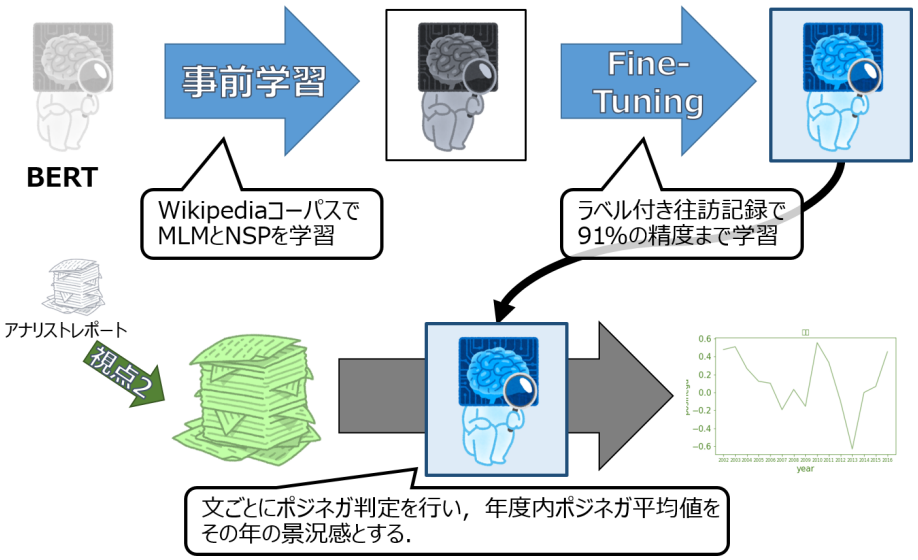


図 3: 景況感推移観測モジュールの概略図

表 3: BERT のパラメーター

パラメーター名	値
最大入力トークン数	398
バッチサイズ	2
BERT 内部の隠れ層の次元数	768
学習率	1e-5
ドロップアウト率	0.1
BERT 内部の隠れ層の活性化関数	GELU[6]
Attention[7] のヘッド数	12
Encoder[8] の層数	12
語彙数	32,000

表 4: 「業績」の関連語

単語	コサイン類似度
収益	0.487
決算	0.479
利益	0.479
ﾊﾞﾘｭｴｰｼｮﾝ	0.460
業績予想	0.453
上期決算	0.437
株価	0.405
受注販売	0.385
受注	0.382
単独決算	0.367

トークン化を行ったものである．SentencePiece の辞書は 32,000 の語彙数で学習されたものである．

本検証における BERT のパラメーターは表 3 のとおりである．

Wikipedia コーパスを用いて事前学習を行った後，キーワード関連文抽出モジュールで得られた文ごとに景況感判定を行うため，Fine-Tuning を行った．BERT の最終層に全結合層を結合し，景況感となる 2 値分類を文ごとに行う．アナリスト往訪記録のうちラベルデータの付与された 1,933 文書を用い，テストデータにおいて 91%程度の分類精度が可能となるまで学習を行った．

4 性能評価

4.1 キーワード関連文抽出モジュール

まず，本モジュールで得られた関連語を表 4 及び表 5 に示す．

直感的な違和感を覚えることはほとんどないと言える．特に，「業績」を入力することによって得られた「バリュエーション」は投資判断において一つの重要な指標を示す単語であり，この単語が抽出されていることは注目に値する．「利益」においても「営業利益」・「経常利益」・「純利益」などの単語が抽出され，類似語を抽出するという目標は達成されたと言える．

次に，「利益」を入力することで抽出された関連語から得られた関連文例を図 4 に示す．

2 文目では，「利益」の単語が含まれていないにも関わ

- 0 年度営業利益実績推定値と [G 製品] 事業の有利子負債削減額が事前ｺﾝｾﾝｻｽを上回った点はｻﾌﾟﾗｲｽﾞ。
- ｸﾞﾙｰﾌﾟﾛｰﾝを圧縮することで、[H 製品] 事業の収益力は更に向上が見込めるとしている。
- 会社の優先順位は、期間損益 (pl) が赤字になってしまうよりもﾌﾘｰｷｬｯｼｭﾌﾛｰをﾌﾟﾗｽにすることで、販売状況がさらに悪化すれば、在庫調整のｽﾋﾟｰﾄﾞはさらに強化され、[I 事業所] の効率は大きく低下するであろう。

図 4: 「利益」の関連文例

表 5: 「利益」の関連語

単語	コサイン類似度
収益	0.629
損益	0.534
営業利益	0.521
経常利益	0.509
純利益	0.494
粗利益	0.485
業績	0.479
最終損益	0.467
税引後利益	0.464
売上	0.460

らず，利益に関係が深いと言える文が抽出されている．しかし 3 文目では，「利益」よりもフリーキャッシュフローを重視しているという内容で，利益に関係が深い文とは言えない．このような文も一部抽出されており，除外したり影響を小さくしたりするシステムを開発する必要があると考えられる．

4.2 景況感推移観測モジュール

まず，キーワード関連文抽出モジュールで得られた各文に対して BERT で景況感判定を行った結果を図 5 に示す．

上から 2 つの文に関しては得られた景況感に関して大きな違和感は無く，正しく推定が行われていることがわかる．しかし 3 番目の文のように，1 文内に景況感が混在している文も存在し，人間によっても景況感を付与することは難しい．また 4 番目の文では「売上未達」のネガティブ情報と，「収益性の向上」のポジティブ情報が両方含まれている．文末にあるポジティブ情報を強調した文であると考えられるが，システムの推定結果はネガティブ情報に注目していると考えられる．5 文目は景況感情報が含まれていない情報である．また，例示はしていないが，過去に発行したアナリストレポートから引用されている場合もあり，これらの文章を景況感判定に用いてしまうと，常に同じ結果を出力してしまうため，景況感のバイアスとなってしまい，変化点の検出が難しくなる．

このように，抽出された関連文に景況感と直接関係のない文やバイアスとなる文章，1 文内に景況感が混在する文などが複数含まれており，これらの文を除外したり，係り受け解析などの手法を用いてさらに細かい文書単位で景況感判定を行う必要があると考えられる．また，4 文目のような文を正しく処理するため，文末に重要な情報を配置する日本語の特性を踏まえたシステムを開発する必要があると考えられる．

次に，景況感を追跡した結果を図 6 に示す．企業 A の純利益の実データと，本システムのキーワードとして「利益」を入力した結果を比較した．純利益の実データは有価証券報告書から取得したものである．

システムから得られた景況感推移と対応する実データの推移に明確な相関を発見することはできなかった．

抽出された文には，会社発表やアナリストの想定と実績との差に関しての記述が含まれている．減益であったとしても想定通りならポジティブな文体で書かれており，増益であったとしても想定を下回るようならばネガティブな文体で書かれている．システムから得られた景況感が大きく振れたタイミングで得られた文を解析することで，これらの差を生み出した要因の抽出やアナリストの深層心理や勘違いを可視化することが可能なシステムとなる可能性がある．

5 結論

本論文では，投資信託運用会社の業務をサポートするため，過去のアナリストレポートを解析し，任意の企業・視点に対する景況感を追跡するシステムを提案した．具体的には，与えたキーワードに基づいてアナリストレポートから関連文を抽出するキーワード関連文抽出モジュールと抽出された関連文の景況感判定を

推定結果:ネガティブ　当社では現在の状況から会社計画の達成はやはり困難、会社計画比 0 億円強の営業利益下ぶれを想定する。

推定結果:ポジティブ　① [A 製品] の拡大・ [B 製品] のｼｪｱｱｯﾌﾟ・ｺｽﾄ削減による収益力強化が続く。

推定結果:ポジティブ　部門利益予想は [C 製品] 事業を 0 億円引き上げ、[D 製品] 部門を 0 億円引き下げた。

推定結果:ネガティブ　会社計画比で、減収増益を予想する主な要因は、[C 製品] 事業において、[E 地域] の [C 製品] の販売減速による売上未達と、[F 国] での値上げの浸透による収益性の向上にある。

推定結果:ポジティブ　【今年の株価維持策】…0 月の説明会は既に案内が届いている。

図 5: BERT における各文の景況感推定結果例

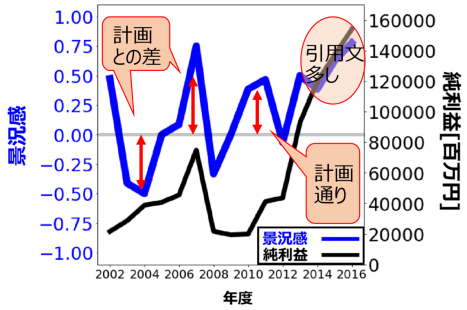


図 6: 企業 A における景況感抽出結果と実データの比較

行って，時期ごとにその推移を観測する景況感推移観測モジュールを開発した．これらモジュールによって，大量に蓄積されたアナリスト往訪記録やアナリストレポートから関心の高い記述を見つけることが容易になり，ファンドマネージャーやアナリストの業務をサポートする当初の目標を一定程度実現できた．

しかし，関連文の抽出精度がキーワードによっては，あまり良くないという課題も残されている．また，得られた景況感と実際の経済指標との明確な相関を確認できず，この点も課題として残されている．一方で，本システムにおける関連度抽出の精度がさらに向上すれば，アナリストの深層心理を可視化できる可能性もあり，更に検証を行っていく必要がある．

参考文献

[1] 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, 2017.

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

[3] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed Representations of Words and Phrases and their Compositionality. In *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc., 2013.

[4] Sato Toshinori. Neologism dictionary based on the language resources on the web for unidic-mecab, 2015.

[5] Taku Kudo and John Richardson. SentencePiece: A simple and language independent subword tokenizer and detokenizer for neural text processing. In *Proceedings of the 2018 Conference on EMNLP: System Demonstrations*, pages 66–71. ACL, 2018.

[6] Dan Hendrycks and Kevin Gimpel. Bridging nonlinearities and stochastic regularizers with gaussian error linear units. *arXiv preprint arXiv:1606.08415*, 2016.

[7] Cicero Nogueira dos Santos Mo Yu Bing Xiang Bowen Zhou Yoshua Bengio Zhouhan Lin, Minwei Feng. A structured self-attentive sentence embedding. *Technical report*, 2017.

[8] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on EMNLP*, pages 1724–1734. Association for Computational Linguistics, 2014.

接触履歴を用いた地方景況感と業種間構造の分析

Analysis of Local Business Confidence indices and Inter-Industry Relations using Contact Histories

坂地泰紀[1*]　蔵本涼太[1]　和泉潔[1]
松島裕康[1]　島田尚[1]　砂川恵太[2]
Hiroki Sakaji[1]　Ryota Kuramoto[1]　Izumi Kiyoshi[1]
Hiroyasu Matsushima[1]　Takashi Shimada[1]　Keita Sunakawa[2]

[1] 東京大学
[1] The University of Tokyo
[2] 沖縄銀行
[2] The Bank of Okinawa

Abstract: 自然言語処理の発展に伴い，膨大なテキストデータから有益な情報を抽出することが可能となりつつある．テキストデータを活用した分析には，その更新頻度の高さやリアルタイム性から金融経済情勢を迅速かつ的確に把握できるという強みがある．そのため，既存指標を補完する速報性の高い指標生成はテキストマイニングとの親和性が高い．本論文では，日次で蓄積されるテキストから景況指標を生成する手法を提案した．そして，実際に沖縄県を対象に景況指標を生成し，既存指標を高い精度で再現していることを確認した．加えて，得られた景況指標から業種間における景況感の連動性を分析した．

1　はじめに

経済・金融において，種々のデータは将来のマーケット動向を予測し，その分析を行うために用いられている．一般的に，マーケット動向の予測手法は扱うデータの差異からテクニカル分析とファンダメンタル分析に大別される．前者は，将来の市場は過去の取引実績の時系列データから予測可能であるという立場の下，過去の取引価格や出来高を入力データとして予測を行う．一方で後者は，テクニカル分析で活用されるデータに加え，該当企業の業績予想など幅広い情報を利用して将来予測を図る手法である．投資家やアナリストは両手法を組み合わせ活用してきたが，その扱い易さからいずれも数値データのみに焦点を当てることが主であった．しかしながら，数値情報以外にもマーケット分析に有益なデータは存在する．その代表例がテキスト情報である．我々は日々様々なテキストから情報を得ている．ファイナンスの分野においても，専門家によって書かれたレポートには経済分析に関する重要な知見が豊富に含まれているであろうし，新聞記事は過去の出来事やその影響を報じている．また，SNS におけるコメントにも人々の経済に対する印象が反映されていると考えることができる．これらテキスト情報を将来動向予測に役立てようとする動きが進んでおり，機械学習を活用した様々な研究が提案されている [Bollen 11]．

地方銀行もまた貴重な金融テキストデータを持っている．例えば，業務日報や決算概況，接触履歴などがある．業務日報や接触履歴の中には，日々の業務の中で行員が感じた些細な変化や気づきが含まれている場合がある．さらに，接触履歴には業種に関する情報も付与されている場合もある．そこで，本研究では，地方景況感と業種間構造の分析を接触履歴を用いて行う．

2　関連研究

機械学習によるインデックス生成に関する研究として，山本らの研究 [山本 16] と余野らの研究 [余野 17] がある．山本らは LSTM を用いて景気ウォッチャー調査を学習し，一例として学習したモデルを用いて月例経済報告 (`http://www5.cao.go.jp/keizai3/getsurei/getsurei-index.html`) を指標化している．ここで，月例経済報告は日本政府が毎月発行する経済レポートである．また，余野らは山本らと同様に景気ウォッチャー調査で学習したモデルを用いて，月例経済報告から指数を生成している．彼らは LDA を用いることにより，個人消費や公共投資，物価などのトピックごとに指標

*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: sakaji@sys.t.u-tokyo.ac.jp

を生成することで，どのファクターが影響しているかを調査できるような手法を提案している．これらの研究では，景気ウォッチャー調査と月例経済報告という類似した文書を利用しているのに対して，我々は接触履歴というテキストを利用している点が異なる．

極性付与に関する研究として，Sakaji et al. の研究 [Sakaji 08] がある．Sakaji et al. は新聞記事から景気動向の根拠となる表現を抽出し，抽出した根拠となる表現に極性を付与する手法を提案している．Turney[Turney 02] は，“excellent” や “poor” などの語を使って，Positive・Negative フレーズを抽出し，それを用いて文書分類を行っている．Pang et al.[Pang 02] は機械学習を用いて極性分類を行う手法を提案している．これらの研究とは異なり，本研究では，地域景況インデックスを生成することを目的にしている．

他にも関連研究として，迫村らの研究 [迫村 13] や山澤の研究 [山澤 18] がある．迫村らは Twitter データを用いて経済動向を分析する研究を行っている．山澤は景気ウォッチャー調査を対象に，単語と景気の現状判断 DI の相関を調べ，さらに主成分分析をすることで，どの単語が景気の現状判断 DI に有効かを調べている．それに対して，我々は主成分分析などの統計的手法ではなく，機械学習を利用することで地域景況インデックスを生成している．

また，文や文書に極性を付与するための辞書を生成する研究として，Ito et al.[Ito 18] の研究がある．Ito et al. は文の極性を判定しながら，辞書も生成する特殊なニューラルネットワークを構築する手法を提案している．加えて，Kaji et al.[Kaji 07] は，人手で作成した手がかり表現リストやパタン，規則から評価文を抽出して極性 (好評，不評) を付与し，各極性中に評価表現 (名詞+格助詞+形容詞) が出現する頻度から評価表現の極性を決定する手法を提案している．Kanayama et al.[Kanayama 06] は，自動的に辞書を拡張する手法を提案している．

3 提案手法

本節では，我々が提案する解析手法の概要について説明する．本手法はまず，Bidirectional long short-term memory(BiLSTM) [Graves 05] を用いて地方景況感を算出する．その後，地方景況感と以下に示すいくつかの手法を用いて，業種間構造の分析を行う．

Step 1: インプットデータとして，景気ウォッチャー調査を用いて BiLSTM モデルを学習する．その後，学習済みモデルを用いて接触履歴に極性を付与し，地方景況感を生成する．

Step 2: 生成した地方景況感に対して，グレンジャー因果分析，インパルス応答分析を用いて業種間構造を分析する．

3.1 地方景況感生成モデル

本節では，地方景況感の生成に用いた BiLSTM モデルについて述べる．我々は，Reccurrent Neural Network(RNN) の一種である LSTM[Hochreiter 97] を機械学習手法として採用した．BiLSTM への入力には，分散表現を用いる．分散表現には 2003 年から 2008 年と 2013 年から 2014 年までのロイターニュース記事と，2014 年から 2017 年までの琉球新報記事を用いる．事前調査でロイターニュース記事だけで分散表現を作ったところ，地域景況インデックス生成において不十分な性能であったため，琉球新報記事を追加した．分散表現生成には gensim[1] を，形態素解析には MeCab[2] を用いた．分散表現は，次元数は 200 次元，5-window で生成した．分散表現に変換した入力に対しては，LSTM($\overrightarrow{LSTM}, \overleftarrow{LSTM}$) を通して $\{\overrightarrow{h_i}\}_i^n$ と $\{\overleftarrow{h_i}\}_i^n$ を得る．その後，$\overleftarrow{h_1}$ と $\overrightarrow{h_n}$ と，0 と 1 で構成されたベクトルに変換した入力 Z を結合する．

$$s = [\overleftarrow{h_1}; \overrightarrow{h_n}; Z] \tag{1}$$

$$t = tanh(W_s \cdot s + b_s) \tag{2}$$

$$u = tanh(W_t \cdot t + b_t) \tag{3}$$

$$v = tanh(W_u \cdot u + b_u) \tag{4}$$

$$Y = W_v \cdot v + b_v \tag{5}$$

ただし，$Z \in \mathbb{R}^z$，$s \in \mathbb{R}^{2m+z}$，$t \in \mathbb{R}^{400}$，$u \in \mathbb{R}^{200}$，$v \in \mathbb{R}^{100}$．z はベクトルの次元数，m は隠れ層のユニット数である．ここで，W_s, W_t, W_u, W_v は重み行列，b_s, b_t, b_u, b_v はバイアスベクトルである．Y は $Y = (y_1, y_2, y_3)$ で構成される出力層である．最後に，出力層の中から最大の値を選び，それに対応する極性を入力に対する出力とする．BiLSTM モデルを図 1 に示す．詳細は，次の論文 [坂地 19] を参照されたい．

3.2 業種間構造分析手法

本節では，業種間構造を分析する手法について述べる．

3.2.1 グレンジャー因果分析

業種間の因果関係をグレンジャーの因果性を検定することで確認する．その具体的な過程は，帰無仮説として業種 A から業種 B へのグレンジャー因果性がないと

[1] https://radimrehurek.com/gensim/
[2] https://taku910.github.io/mecab/

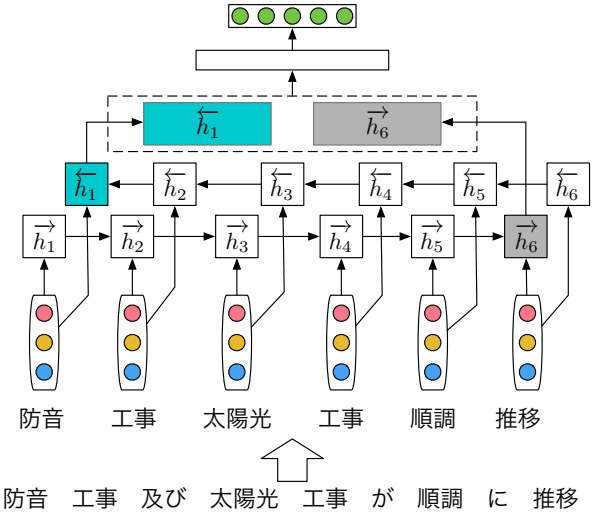


図 1: Bidirectional LSTM を用いての極性判定 [坂地 19]

仮定し，同仮説を棄却するか否かをその p 値と有意水準との比較により判断するというものである．本研究では，グレンジャー因果を多変量へ拡張した偏グレンジャー因果分析 [Guo 08] を採用した．偏グレンジャー因果分析は外生要因や潜在変数の影響を減じることが可能であり，多変量データ解析において有用である．

3.2.2 インパルス応答分析

グレンジャー因果検定は時系列データ間の因果性を判定する手法であるが，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，その関係性を定量的に捉える手段として用いられる．ある変数における変動が他の変数の変動に寄与する度合いを分析することで，定量的な解析が可能となるのである．本論文では，複数業種の景況指数が与えられたときに，ある業種での変動が他業種にどのような影響を及ぼすのか定量的評価を行うため，直交化インパルス応答関数 [Sims 80] を活用し分析する．

4 評価実験

本研究では，沖縄県に着目し，沖縄銀行の接触履歴を用いて地方景況感と業種間構造を分析する．

4.1 地方景況感生成実験に用いるデータ

我々は地方景況感生成実験に以下のデータを用いる．

表 1: 接触履歴の例

他行住宅ローン借換推進で訪問するが不在。名刺、チラシ投函する。
修学旅行も引続き順調であり来期も期待できる。新ホテル棟建設についてであるが投資金額が当初予定額より大幅に上回り、再度計画予定である。一応は 9 月までを目処としているが、建築確認等を考えると厳しいかもしれない。今後の展開を考え、＜組織＞内等にホテル等の売却案件があったら運営及び買取を行いたいので是非紹介をお願いしたい。
＜組織＞社長面談　現況確認　台風の影響で前期は売上減少　健康食＜組織＞による売上増加を見込むが動きなし　同氏の土地収用に伴う入金で借入金返済を見込む
キャッシュカード暗証番号変更でご来店。定期商品や保険を少しご案内しました。老後のことが不安との事なので時間があるときに検討する。

琉球新報　沖縄県の地方新聞紙として，琉球新報がある．本研究では，2014 年から 2017 年の琉球新報を用いる．

接触履歴　接触履歴は，行員が顧客と何か知らのやり取りを行った際に記録されるデータであり，様々なことが記述されている．表 1 にその例を示す．

4.2 地方景況感生成実験設定

我々は，接触履歴，もしくは，琉球新報を用いて地方景況感を生成する実験を行う．琉球新報においては，文，段落，記事全体のそれぞれを用いて生成実験を行う．これは，新聞記事中のどの箇所が地方景況感を表しているのかを調べるためである．正解データとして，おきぎん経済研究所が発行している企業動向調査[3]を用いる．

5 結果と考察

表 2 に地方景況感生成実験の結果を示す．表 2 より，対象テキストの範囲変化による効果検証実験において，景気に関連のある記事・段落・文に大きな差異はないことが分かった．それに対して，接触履歴を用いた地方景況感生成は非常に高い精度で再現できている．その結果を図 2 を示す．

[3] http://www.okigin-ei.co.jp/report_DI.html

表 2: 地方景況感生成実験の結果

	相関係数
琉球新報：文単位	0.301
琉球新報：段落単位	0.275
琉球新報：記事単位（景気）	0.300
琉球新報：記事単位（全体）	0.175
接触履歴	0.856

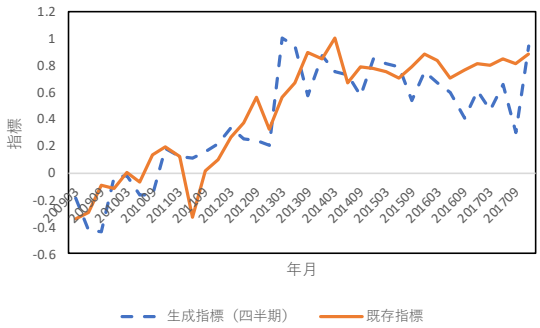


図 2: 接触履歴により生成した景況指標

次に，業種間分析の結果を図 3 と図 4 に示す．それぞれ，接触履歴から生成した地方景況感を用いて分析した結果である．沖縄県において，業種の影響力係数を大きい順に列挙すると，商業（小売業と卸売業），建設業，サービス，製造業，医療・保健・社会保障・介護，不動産業の並びになる．そして，この並びは景況感変動の分析結果と結びつけて考えることが可能である．上位 3 つの業種，すなわち，商業（小売業と卸売業）と建設業，サービスは，図 3 の上流に位置する業種と対応している．また，図 4 に示すように，インパルス応答分析においても他業種に対する大きな変動要因になっている．そのような業種の景況感が向上することで，他業種に生産を喚起し，それに伴い他業種の景況感も向上する．この流れは理にかなったものであり，他業種への景況変動要因となる業種は生産波及効果の高い業種に対応することが確認された．

6 まとめ

本論文では，日次で蓄積されるテキストから地方景況感を生成する手法を提示した．そして，実際に沖縄県を対象に地方景況感を作成し，既存指標を高い精度で再現していることを確認した．加えて、得られた地方景況感から業種間における景況感の連動性を明らかにした．

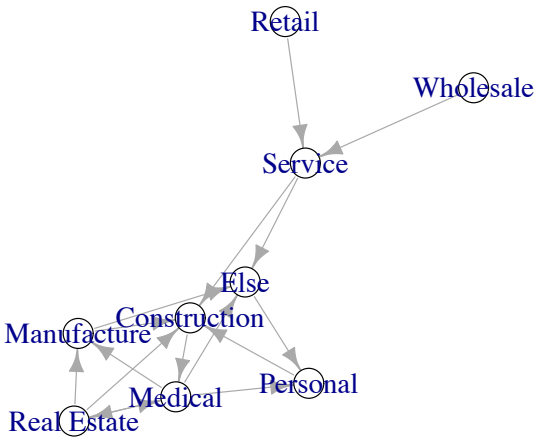


図 3: 偏グレンジャー因果分析による業種間因果の可視化 [Sakaji 19]

謝辞

本研究は，三井情報株式会社と株式会社沖縄銀行，琉球新報社の協力により行うことができた．三社に感謝いたします．

参考文献

[Bollen 11] Bollen, J., Mao, H., and Zeng, X.: Twitter mood predicts the stock market, *Journal of computational science*, Vol. 2, No. 1, pp. 1–8 (2011)

[Graves 05] Graves, A. and Schmidhuber, J.: Framewise phoneme classification with bidirectional LSTM and other neural network architectures, *Neural Networks*, Vol. 18, No. 5, pp. 602–610 (2005)

[Guo 08] Guo, S., Seth, A., Kendrick, K., Zhou, C., and Feng, J.: Partial Granger causality?Eliminating exogenous inputs and latent variables, *Journal of Neuroscience Methods*, Vol. 172, No. 1, pp. 79–93 (2008)

[Hochreiter 97] Hochreiter, S. and Schmidhuber, J.: Long short-term memory, *Neural computation*, Vol. 9, No. 8, pp. 1735–1780 (1997)

[Ito 18] Ito, T., Sakaji, H., Izumi, K., Tsubouchi, K., and Yamashita, T.: GINN: gradient interpretable neural networks for visualizing financial texts, *International Journal of Data Science and Analytics* (2018)

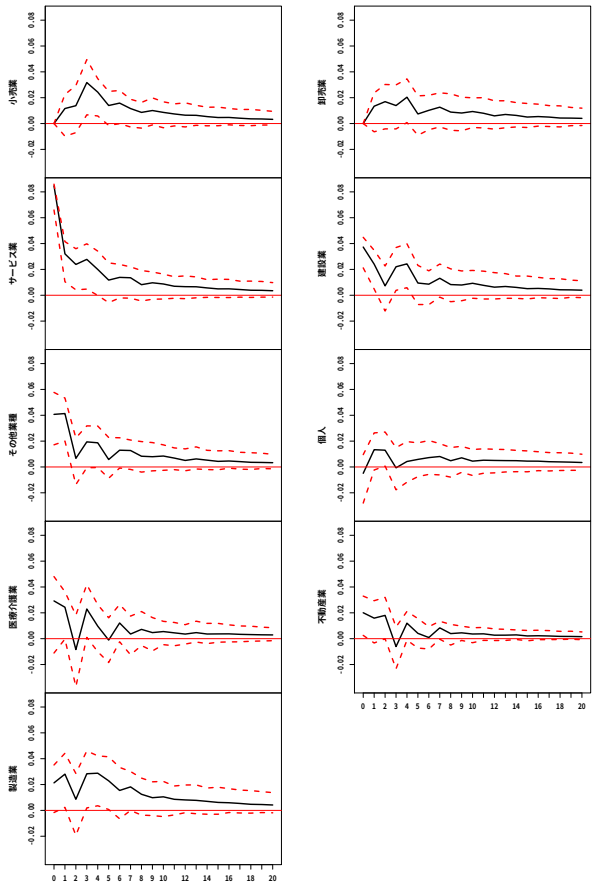


図 4: インパルス応答関数（サービス業）

[Kaji 07] Kaji, N. and Kitsuregawa, M.: Building Lexicon for Sentiment Analysis from Massive Collection of HTML Documents, in *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL*, pp. 1075–1083 (2007)

[Kanayama 06] Kanayama, H. and Nasukawa, T.: Fully Automatic Lexicon Expansion for Domain-oriented Sentiment Analysis, in *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, EMNLP '06, pp. 355–363, Stroudsburg, PA, USA (2006), Association for Computational Linguistics

[Pang 02] Pang, B., Lee, L., and Vaithyanathan, S.: Thumbs Up?: Sentiment Classification Using Machine Learning Techniques, in *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing - Volume 10*, EMNLP '02, pp. 79–86, Stroudsburg, PA, USA (2002), Association for Computational Linguistics

[Sakaji 08] Sakaji, H., Sakai, H., and Masuyama, S.: Automatic Extraction of Basis Expressions That Indicate Economic Trends, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 977–984 (2008)

[Sakaji 19] Sakaji, H., Kuramoto, R., Matsushima, H., Izumi, K., Shimada, T., and Sunakawa, K.: Financial Text Data Analytics Framework for Business Confidence Indices and Inter-Industry Relations, in *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pp. 40–46, Macao, China (2019)

[Sims 80] Sims, C. A.: Macroeconomics and reality, *Econometrica: journal of the Econometric Society*, pp. 1–48 (1980)

[Turney 02] Turney, P. D.: Thumbs Up or Thumbs Down?: Semantic Orientation Applied to Unsupervised Classification of Reviews, in *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, ACL '02, pp. 417–424, Stroudsburg, PA, USA (2002), Association for Computational Linguistics

[坂地 19] 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)

[山本 16] 山本 裕樹, 松尾 豊：景気ウォッチャー調査の深層学習を用いた金融レポートの指数化, 2016 年度人工知能学会全国大会 (2016)

[山澤 18] 山澤 成康：計量テキスト分析による景気判断－コーディングルールや主成分を使った時系列分析－, ESRI Discussion Paper Series (2018)

[迫村 13] 迫村 光秋, 和泉 潔：twitter テキストマイニングによる経済動向分析, 第 9 回人工知能学会ファイナンスにおける人工知能応用研究会 (2013)

[余野 17] 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, 2017 年度人工知能学会全国大会 (2017)

外国為替市場におけるポジション管理戦略分類に基づく将来価格形成メカニズム解明

Clarification of mechanism underlying price formation process in a forex market on the basis of clustering of position management strategies

末重拓己 [1]　Didier Sornette[2]　高安秀樹 [3,4]　高安美佐子 [1,4]

Takumi Sueshige[1], Didier Sornette[2], Hideki Takayasu[3,4], and Misako Takayasu[1,4]

[1] 東京工業大学情報理工学院
[1]Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology
[2]ETH Zurich, Department of Management
[3]Sony Computer Science Laboratories
[4] 東京工業大学科学技術創成研究院
[4] Institute of Innovative Research, Tokyo Institute of Technology

Abstract: Individual trading analysis has become the focus of recent attention in the field of market microstructure. Several research works reveal how trading strategies of specific traders or banks are affected by historical market information. However, there is little research revealing how such microscopic trading strategies recursively affect macroscopic future market information. Using the high-granular dataset including the trading behavior of specific banks in a U.S dollar (USD) against Japanese yen (JPY) market, here we demonstrate management method of positions, defined as the numbers of units of USD banks bought or sold against JPY, can be clearly clustered into two simple strategies. We then find the strong relationship between future-market price movements and these two position management strategies, and this relationship even allows a prior prediction of market prices fifteen minutes ahead.

1 導入

2010 年頃より，社会系の分野においてビッグデータ解析の有用性が認知され始め，至る所で積極的にデータを収集しようという試みがなされている[1-2]. 社会系のうち特に金融市場では，電子技術の発達も相まって，他の分野に先立って膨大なデータが蓄積されてきている[3-6]．金融市場のデータで近年注目を集めているのが, 個々のトレーダや銀行の ID が付随されたデータである[3-7]．トレーダないし銀行 ID はデータ期間で固定されているため，データがカバーしている範囲においてトレーダや銀行がどのタイミングでどれくらいの注文量を約定させているのかという履歴をトラッキングすることが可能となる．このようなデータ特性から，取引に至るまでのトレーダ・銀行のミクロな戦略解明が期待されており，事実，先行研究[3-7]では，過去の取引価格・取引量やニュースなどのマクロ変数により条件付けることで戦略の同質性や異質性が明らかにされ始めている．

しかし，トレーダや銀行ごとのミクロな戦略解析が進みつつあるものの，明らかになったミクロ戦略がマクロな取引価格の性質にどのような影響を与えるかという研究はほとんどされていない．そもそもミクロ戦略の解析が始められたモチベーションの一つに，金融市場の戦略多様性と現実の生態系における種の多様性のアナロジーから，金融市場の理解に生態学的アプローチを援用しようという発想（market ecology）がある[8-9]．端的に言えば，market ecology では，採用される戦略の時間発展や，戦略の多様性を通して，個々のミクロな戦略からマクロな取引価格の統計則解明が期待されている．この発想を実現させるためには，金融市場で観測されるミクロ戦略の詳細解析は当然必要であるが，ミクロ戦略がどのような条件下でマクロな取引価格の統計則と繋がりを見せるかといった，観測スケールを跨ぐ解析が難関であると同時に最も重要な研究領域である．

そこで，本研究では，論文[10]で明らかになった，Electronic Broking System (EBS)社提供のトレーダ・銀

行 ID 付き高頻度データを用いて, ポジションと呼ばれる，日本円と米ドルのネットの累積取引量に着目し，銀行がポジションをどのように管理しているかという点からミクロ戦略を 2 種類のクラスタに分類する方法を紹介する．そして各クラスタのポジション総量の偏りが十分大きくなったとき，ポジション管理戦略が強く将来価格を歪めることを示し，実際に 15 分程度将来の価格予測が可能となることを報告する．

2. 解析データと解析手法

2.1 解析データと用語の定義

本研究では，2016 年 6 月 5 日から 10 日におけるドル円インターバンク市場における millisecond ごとに記録されたトレーディングログを用いて解析を行う．先に述べたように，このデータでは EBS 市場に発注される注文にトレーダ・銀行 ID が付与されているため, 1 週間の取引履歴をトレーダ・銀行ごとにトラッキングすることが可能である．なお，本研究ではトレーダを所属する銀行ごとに合算して解析を行うため, 以下では銀行 ID のみ用いる．解析対象週における注文価格の最小単位を 0.001 円とし，これを tpip と呼ぶ．また注文量の最小単位は 1M$であり，これを最小単位とし, 最小単位の倍数を本数と呼ぶ．

解析手法の詳細に入る前に，ポジション関連の用語をここで定義する．経済ニュースなどでドル円，ユーロ円市場と呼称されることわかるように，為替通貨の取引は必ず 2 通貨間で行われる．この 2 通貨のうち，先頭にある通貨（例えばドル円の場合はドル）を基軸通貨，後に記載されている通貨（ドル円の場合は円）を参照通貨と呼ぶ．為替市場における通貨の取引価格は，参照通貨購入のためには，いくらの基軸通貨売却が必要となるのかという 2 通貨間の交換レートを表しているといえる．ドル円市場で 1$=110 円というレートは，参照通貨（円）を 110 円購入するためには，基軸通貨（ドル）を 1$の売却が必要であることを意味している．このとき，トレーダが参照通貨を購入するときの取引をロングコントラクト，売却するときの取引をショートコントラクトと呼ぶ．本研究ではコントラクトの単位を注文量で定義し，ロングコントラクトを正，ショートコントラクトを負として計算する．そのため， 1 週間を通してロングコントラクトとショートコントラクトの両方を締結しているとき累積コントラクトが正であればロングポジション，負であればショートポジションと呼ぶ.

2.2 解析手法

ここでは，2.1 で定義したポジションを用いて，ポジション管理をクラスタリングする手法を説明する．ポジション管理を特徴づける変数として，過去平均ポジションを導入する．そのために，まず数式を用いてポジション定義しておく．

$$V^i(t) \equiv \sum_{k=1}^{N^i(t)} v^i\left(T_k^i\right).$$

ここで，上付きのiは銀行iのインデックス，$N^i(t)$は時刻tまでに取引した回数，T_k^iはk回目に取引した時刻，$v^i(t)$は時刻tにおける取引量（ロングコントラクトは正，ショートコントラクトは負）である．このとき，κ-tick に渡る過去平均ポジションは

$$\bar{V}^i(t) = \frac{1}{\kappa}\sum_{k=0}^{\kappa-1} V^i\left(T^i_{N^i(t)-k}\right),$$

と定義する．なお，ポジションは整数値しかとらないため，過去平均ポジションを四捨五入した値を$[\bar{V}^i(t)]$とし，以下の解析ではこの値を用いる．

3. 解析結果

3.1 ポジション管理の戦略分類

図 1(a)に，取引頻度 2 位と 4 位の銀行について，現在のポジションと過去平均ポジションの差分（以下ポジション差分）$V^i(t) - [\bar{V}^i(t)]$の符号で条件付けしたときの，次の取引時刻における取引量がロングコントラクトである確率を示した．図より，取引頻度 2 位の銀行はポジション差分が正のときは 80%程度の確率で次の取引でロングコントラクトを選択しており，ポジションを積み上げていく傾向が見て取れる．一方，取引頻度 4 位の銀行はポジション差分が正のときは 75%程度の確率で次の取引にショートコントラクトを選択しており，ポジションの偏りを打ち消す方向に取引を行っている傾向が見て取れる．

これら関係式が 2 行に特有な性質ではなく，複数銀行のポジション管理戦略として観測されることを示すため，取引頻度の上位 30 行について$V^i(t) - [\bar{V}^i(t)]$と$v^i(t)$の関係を図(b)に示した．図の色分けは, gnuplot の Levenberg-Marquart アルゴリズムを用いて線形フィッティングを行った結果，フィッティングエラーを考慮しても傾きが正である場合には橙色，傾きが負である場合には緑色とした．この分類により, 30 行のうち 27 行について, 図 1(a)と同様の結果が成り立っていることがわかり，ポジション差分と将来の取引量の相関が 2 行に特有の統計則ではない

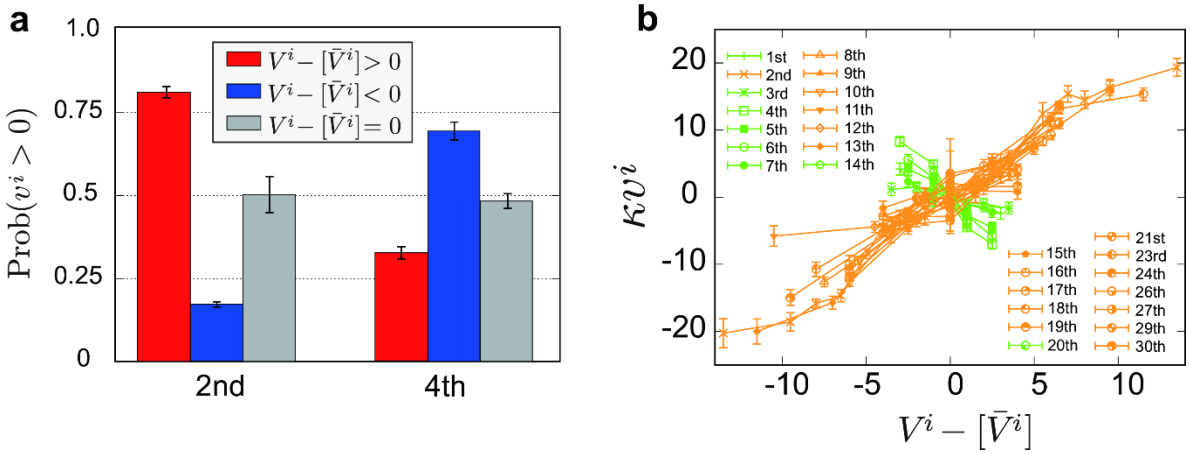


図 1 ポジション管理の統計則[10]．(a)取引頻度 2 位と 4 位の銀行について，ポジション差分で条件付けたときの，次の取引時刻の取引がロングコントラクトである条件付き確率．(b)取引頻度上位 27 行のポジション差分と次の時刻の取引量の線形関係．

ことがわかる．本稿では，ポジション差分と取引量の関係が正の線形関係である銀行を，経済的需給に基づいて取引を行っている economic-motivated (EM) 銀行，負の線形関係である銀行を，裁定取引を狙いとして取引を行っている arbitrage-motivated (AM)銀行と呼ぶことにする．なお，解析対象とした銀行 30 行が 1 週間の取引回数に占める割合は 56.1%であり，全体の取引の過半数を占める銀行の統計則を明らかにしたことになる．

3.2 ポジションの偏りを用いた将来価格予測

ここでは, 3.1 で明らかにしたポジション管理戦略の分類を用いて，過去τ分のデータから$\tau/2$分将来の価格予測を行う．そのために，以下で示した，分類した 2 種類の戦略を持つ EM, AM 銀行のポジションの偏りの総和を予測シグナルとして定義した．

$$s^C(t) \equiv \sum_{i \in C} \sum_{k}^{\tau} \frac{1}{\tau} \{V^i(t-k) - [\bar{V}^i(t-k)]\}.$$

ここで，$C \equiv \{\{\mathrm{EM}\}, \{\mathrm{AM}\}\}$でありそれぞれ EM 銀行，AM 銀行のクラスタ，iは各クラスタに属する銀行のインデックス，τは過去のポジション差分の平均をとる look back 期間(単位：分)を意味している．なお，look back 期間内の平均を計算するにあたり，分刻みのポジション差分データを用いた．

図 2 に， EM 銀行(a)，AM 銀行(b)全体のポジション差分の偏りが閾値γ^Cを超えたときをシグナルとし，シグナル検出時刻から$\tau/2$分後の将来価格の分布を示した．図の横軸は，シグナル検出時刻の価格を基準値としたときの$\tau/2$分後の価格変化を表している．赤線が閾値γ^Cよりも正方向に大きな偏りが生じたときの価格変化分布，赤線は閾値$-\gamma^C$よりも負方向に大きな偏りが生じたときの価格変化分布，灰色は閾値を超えないときの価格変化分布を表している．なお，この図では$(\tau, \gamma^{\mathrm{EM}}, \gamma^{\mathrm{AM}}) = (30,8,6)$としている．

図より，ミクロな銀行単位のポジション管理戦略がマクロな取引価格分布に強い影響を及ぼしていることがわかる．具体的には, EM 銀行のポジションの偏りをシグナルとして用いた図 2(a)では，EM 銀行全体で正のポジションの偏りがある場合，将来価格分布が正の方向に歪み，反対に負のポジションの偏りがある場合，将来価格が負の方向に歪んでいる．これは，EM 銀行が正（負）方向にポジションが偏っている場合，図 1 の関係式より将来の取引もロングコントラクト（ショートコントラクト）が生じやすい．そして, EM 銀行が全体としてロングコントラクト（ショートコントラクト）を選択すると，将来価格が上昇（下落）するため，EM 銀行全体のポジションの偏りが将来価格を歪めたものと考えられる．一方，AM 銀行全体として正のポジションの偏りがある場合，将来価格が負の方向に歪み，反対に負のポジションの偏りがある場合には将来価格が正の方向に歪んでいる．この関係は EM 銀行の場合と同様に考察することができる．図 1 より AM 銀行はポジションの偏りを打ち消す方向に次の取引を選択する傾向があった．このポジション管理戦略より，AM 銀行全体として正のポジションの偏りがある場合，次の取引にショートコントラクト（ロングコントラクト）を選択する確率が高まり，将来価格が下落（上昇）するため，AM 銀行全体のポジションの偏りが将来価格を歪めたものと考えられる．

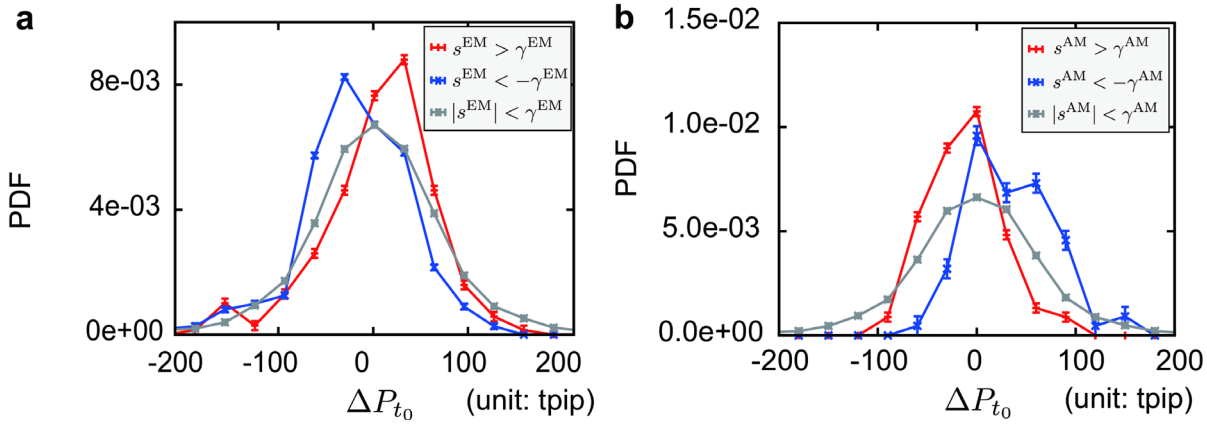


図 2 EM 銀行(a)と AM 銀行(b)のポジション差分の偏りで条件付けしたときの将来価格変化分布[10].

4. まとめと今後の課題

本稿では，銀行のポジション管理戦略に着目し，ポジション管理の方法を，(i)ポジションを大きく積み増していくよう取引する銀行と，(ii)ポジションの偏りを減じる方向に取引する銀行の 2 種類に分類した．そして，分類した銀行全体のポジションが強く偏ったとき，銀行のポジション管理戦略により将来価格が歪められることを明らかにした．この研究により金融市場の銀行単位のミクロ戦略と，マクロな取引価格の関係の一部が明らかになり，金融市場の理解を個々のトレーダないし銀行レベルの取引戦略から解明していく market ecology の研究を推し進めたと考える．より詳細な解析や過去平均ポジションの解釈については論文[10]を参照されたい．

本研究で明らかにした銀行のポジション戦略分類と将来価格予測の関係を実務的に使用するためには解決すべき問題が大きく 2 つある．一つ目は，データ自体の問題である．金融市場の公開データから各銀行のポジションの偏りを推定することはできない．そのため，公開データからポジションの偏りを推定するような新たな研究が必要となる．二つ目は，パラメータ推定の問題である．将来価格予測のためには，本研究で導入した$(\tau, \gamma^{\mathrm{EM}}, \gamma^{\mathrm{AM}})$のパラメータを事前に推定する必要がある．論文[10]で示した通り，将来価格分布は各パラメータの摂動に対してある程度ロバストであるものの，リアルタイムトレーディングの目的のためには，例えば前週のデータを使ってこれらパラメータを推定し，翌週のトレーディングに用いるという推定プロセスが必要となる．現状複数週連続したデータがないため厳密な推定プロセスの検証は行えなかったが，データが入手可能になり次第推定プロセスの問題に取り組むことを考えている．

謝辞

本研究を進めるにあたり，金澤輝代士氏（筑波大），木島俊哉氏（三菱 UFJ），杉浦大地氏（三菱 UFJ），北野尚吾氏（三菱 UFJ）との議論が参考になりました．ここで謝辞させていただきます．また，本研究は学術振興会特別研究員研究奨励費 (No. 17J10781)，基盤研究(B) (No. 26310207)，戦略的国際共同研究プログラム（SICORP），MUFG AI 金融市場解析共同研究講座，the Eidgenössische Technische Hochschule grant 0-20067-14 on “Quantifying Financial (in-)stability and self-amplification of crashes via reflectivity”から援助されていること明記させていただきます．

参考文献

[1] Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. MIS quarterly, 36(4).

[2] Lynch, C. (2008). Big data: How do your data grow? Nature, 455(7209), 28.

[3] Lillo, F., Micciché, S., Tumminello, M., Piilo, J., & Mantegna, R. N. (2015). How news affects the trading behaviour of different categories of investors in a financial market. Quantitative Finance, 15(2), 213-229.

[4] Tumminello, M., Lillo, F., Piilo, J., & Mantegna, R. N. (2012). Identification of clusters of investors from their real trading activity in a financial market. New Journal of Physics, 14(1), 013041.

[5] Musciotto, F., Marotta, L., Piilo, J., & Mantegna, R. N. (2018). Long-term ecology of investors in a financial market. Palgrave Communications, 4(1), 92.

[6] Kanazawa, K., Sueshige, T., Takayasu, H., & Takayasu, M. (2018). Derivation of the Boltzmann equation for financial

Brownian motion: Direct observation of the collective motion of high-frequency traders. Physical review letters, 120(13), 138301.

[７] Sueshige, T., Kanazawa, K., Takayasu, H., & Takayasu, M. (2018). Ecology of trading strategies in a forex market for limit and market orders. PloS one, 13(12), e0208332.

[８] Farmer, J. D., & Skouras, S. (2013). An ecological perspective on the future of computer trading. Quantitative Finance, 13(3), 325-346.

[９] Farmer, J. D. (2002). Market force, ecology and evolution. Industrial and Corporate Change, 11(5), 895-953.

[１０] Sueshige, T., Sornette, D., Takayasu, H., & Takayasu, M. (2019). Classification of position management strategies at the order-book level and their influences on future market-price formation. PLOS ONE, 14(8), e0220645..

リターンに基づくクラスタリングを用いた投資信託スタイル分析

Mutual fund style analysis using clustering based on return

副島聡一郎 [1*] 松井 藤五郎 [2] 犬塚 信博 [1] 武藤 敦子 [1] 森山 甲一 [1]
Soichiro Soejima[1] Tohgoroh Matsui[2] Nobuhiro Inuzuka[1] Atsuko Mutoh[1] Koichi Moriyama[1]

[1] 名古屋工業大学
[1] Nagoya Institute of Technology
[2] 中部大学
[2] Chubu University

Abstract: This paper proposes a clusterling of investment trusts based on returns in order to classify mutual funds into actual funds style groups. In the conventional method, the investment trusts are clustered based on the similarity of the investees. However, since the investees cannot be confirmed by anyone other than the management company, they cannot be used by anyone other than the management company. Analyze the investment trust using the proposed method and confirm that the mutual fund styles is included in the similar investment trust or included in the cluster.

1 はじめに

近年、大量のデータから知識発見、問題解決をすることが重要とされており、データマイニング が注目されている。その手法の一つとしてクラスタリングが存在し、データに対して一定の基準からグループ分けを行いグループごとに特徴や性質を発見することができ、金融商品の分類にも活用されている。

投資信託 (ファンド) とは、販売会社を通じて複数の投資家から集めた資金を、資産運用の専門家が株や債権などの金融資産、あるいは不動産などに投資するように指示し、運用成果が投資家それぞれの投資額に応じて分配される仕組みの金融商品である。

ファンドの運用方針はインデックス型とアクティブ型に分けることができる。インデックス型とは株式相場の指標に用いられるインデックス (日経平均株価や東証株価指数など) に連動する運用方針を持つファンドであり、アクティブ型とはインデックスを超過することを目標としているファンドである。ファンドごとに目標としているインデックスは異なり、目標としているインデックスのことをベンチマークという。

ファンドの中には公表している運用方針と実際の運用方針が異なる場合がある。そのため投資家は公表している運用方針と実際の運用方針が一致しているかを知りたく、運用会社は公表している運用方針と実際の運用方針と一致していることを示すことで投資家から信頼を得ることができる。

そこで本研究では、リターンに基づいてファンドをクラスタリングすることによってファンドのスタイルを分析する方法を提案する。

*連絡先： (名古屋工業大学情報工学科犬塚研究室)
(名古屋 市昭和区御器所町)
E-mail:s.soejima.417@nitech.jp

2 従来研究

米田ら [1] はファンドの全銘柄の投資割合に基づいた投資先類似度を定義し、実際の運用方針に基づくファンドのグループ分けを目的にファンドをクラスタリングする方法を提案した。

米田らの研究では投資先の情報を用いているが、投資先の情報は運用会社しかわからないため実用的ではないという問題がある。さらに、米田らの研究では実際の運用方針に基づくファンドのグループ分けを目的にしているが、米田らの手法を用いて得たクラスタは運用方針ではなくベンチマークごとに分かれている。

松井ら [2] はリターンの時系列に基づく動的クラスタリングを用いて投資信託を分析する手法を提案し、TOPIX 連動型インデックス・ファンドを分析した。

松井らの手法は同一ベンチマークのインデックス型投資信託、すなわち、同一の運用方針を取っている投資信託のみを対象としていた。本研究では、この手法

を様々な運用方針の投資信託に適用する方法に拡張し、ファンドのスタイル分析を行う。

3 提案手法

本研究では同一のベンチマークを対象とした、月次リターンの類似度に基づいたファンドのクラスタリングを提案する。月次リターンは 3ヶ月の月次リターンデータを 3 次元データとしてクラスタリングを行う。またクラスタリング は k-means 法を用いる。

米田らの手法では投資先の情報は運用会社しかわからないという問題点があった。そこで本研究では外部から観測可能な月次リターンの時系列データを用いることで従来研究の問題を解決する。

ファンドはベンチマークとしているインデックスを目標としているため、結果的にインデックスに類似したリターンになる。従って、ベンチマークが異なるファンドが混在するデータでクラスタリング した場合、ベンチマークが同じもの同士は運用方針に関わらず、ベンチマークが異なるファンド同士に比べて相対的にリターンが似るためベンチマークごとにクラスタが形成されると考えられる。

本研究ではベンチマークごとにクラスタリングを行うことによりこの問題を解決する。

4 予備実験

本章ではベンチマークが異なるファンドが混在するデータを用いてクラスタリングを行うことにより、ベンチマークが異なるファンドが混在しているデータでは運用方針ごとにクラスタが形成されないことを示す。提案手法とはベンチマークが同一のファンドではないという点以外は同じ手法を用いて実験を行う。

4.1 実験方法

用いたデータは同一の運用会社の国内株式に投資している 67 のファンドの 2015 年 10 月から 12 月の月次リターンデータである。月次リターンデータを図 1 に示す。これらのファンドの中には運用方針がインデックス型であるファンドとアクティブ型であるファンドが混在している。またベンチマークとしている指標は TOPIX、日経 225、JASDAQ などがある。名前に TOPIX を意味する言葉を含むファンドを TOPIX ファンド群と呼ぶ。このファンド群の中には 5 つのファンドが含まれる。また日経 225 を意味する言葉を含むファンドを日経 225 ファンド群とする。このファンド群の中には 5 つのファンドが含まれる。TOPIX ファンド群と日経 225 ファンド群に含まれるファンドはいずれもインデックス型のファンドであり、ベンチマークは異なるが運用方針は同じである。

10 月、11 月、12 月の月次リターンを三次元データとして、クラスタリングを行った。分割数は 6 とした。

4.2 実験結果

結果を図 2 に示す。TOPIX ファンド群を構成する 5 つのファンドは全て同じクラスタに含まれた。このクラスタは 23 つのファンドで構成されており、そのうち 5 つが TOPIX ファンド群のファンドである。また日経 225 ファンド群を構成する 5 つのファンドは TOPIX ファンド群とは異なるクラスタですべて同じクラスタに含まれていた。このクラスタは 18 のファンドで構成されており、そのうち 5 つが日経 225 ファンド群のファンドである。

67 のファンドの 10 月、11 月、12 月の月次リターンをクラスタの色ごとに分けたものを図 3 に示す。このように同じクラスタの月次リターンは似たような値動きをしている。

4.3 考察

リターンを用いてクラスタリングをした結果、TOPIX ファンド群や日経 225 ファンド群のようにベンチマークごとにクラスタが形成された。これはファンドごとにベンチマークを目標にしているので、ベンチマークとしているインデックスと似たような動きになる。同じベンチマーク同士はインデックス型であろうがアクティブ型であろうが、他の指標をベンチマークとしているファンドに比べて相対的にリターンの値動きが似るため、ベンチマークごとにクラスタが形成されたと考えられる。例えば TOPIX をベンチマークとしているファンドはインデックス型でもアクティブ型でも TOPIX に連動しようとするか、TOPIX を超過しようとする。つまり TOPIX を基準に考えているため、TOPIX の値動きに近づく。また日経 225 をベンチマークとしているファンドは同様の理由で日経 225 の値動きに近づく。よってベンチマークが異なるファンドが混在したままクラスタリングを行うと、ベンチマークごとにクラスタが分かれると考えられる。

5 実験

予備実験より、ベンチマークが混在したファンド群に対してクラスタリングを行うとベンチマークごとにクラスタが形成されることが確認された。そこで本章ではベンチマークが同じものだけでクラスタリングを

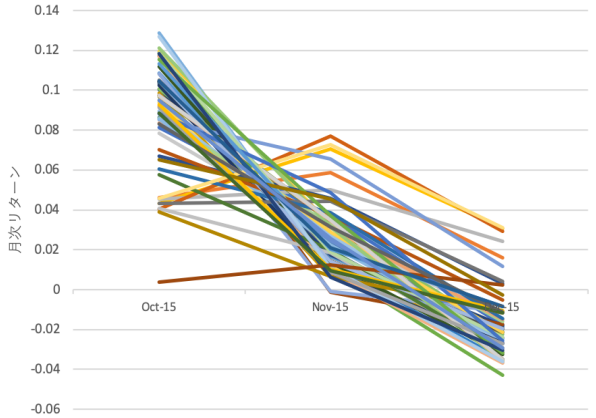


図 1: 67 のファンドの 10 月、11 月、12 月の月次リターン

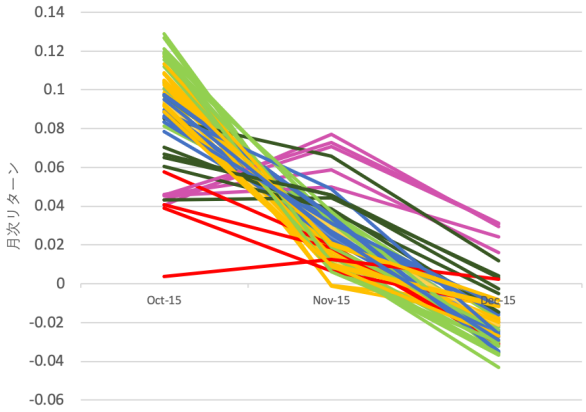


図 3: 67 のファンドの 10 月、11 月、12 月の月次リターンで、色はクラスタ番号を表す

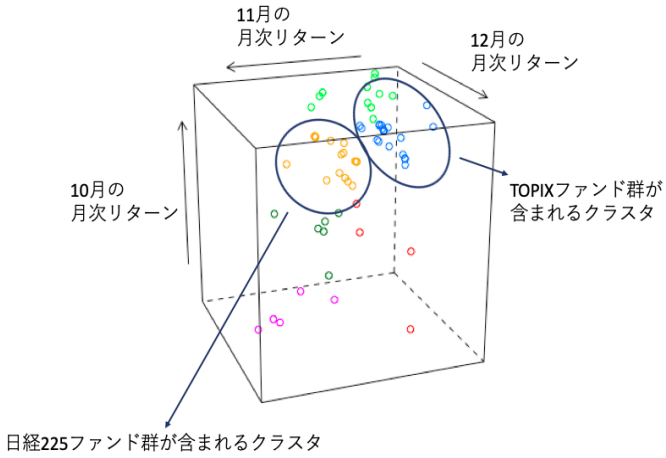


図 2: 67 のファンドを用いたクラスタリング結果

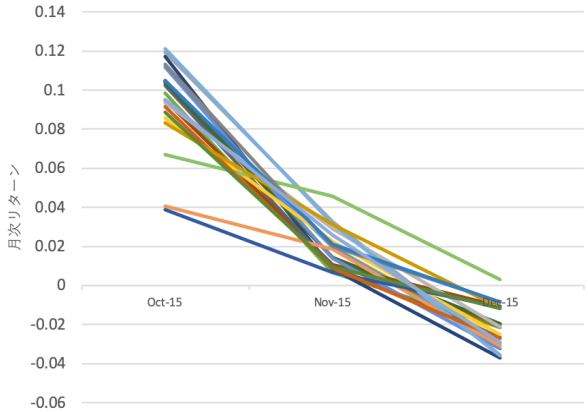


図 4: TOPIX をベンチマークとしている 25 のファンドの 10 月、11 月、12 月の月次リターン

行うことにより、運用方針ごとにクラスタが形成されることを示す。

5.1 実験方法

用いたデータはある同一の運用会社であり、目論見書で TOPIX をベンチマークとしている 25 のファンドの 2015 年 10 月から 12 月の月次リターンデータである。月次リターンデータを図 4 に示す。10 月、11 月、12 月の月次リターンを三次元データとして、クラスタリングを行った。分割数は 4 とした。

5.2 実験結果

結果は図 5 のようになった。2 つのファンドで構成されている孤立しているクラスタが確認できた。これらのファンドは目論見書では運用方針をアクティブとしているファンドであり、ファンド名にアクティブや成長株など運用方針を意味する言葉を明記している。

最も大きな 17 つのファンドで構成されているクラスタには目論見書で運用方針をインデックスとしているファンドが全て含まれた。

25 のファンドの 10 月、11 月、12 月の月次リターンをクラスタの色ごとに分けたものを図 6 に示す。このように同じクラスタの月次リターンは似たような値動きをしている。

5.3 考察

ファンド名にアクティブを意味する言葉や、成長株を意味する言葉を含むファンドは運用方針に偽りがないと考えられる。そのため、これら二つのファンドで構成されているクラスタは運用方針がアクティブであるグループだと考えらえる。最も大きな 17 つのファンド

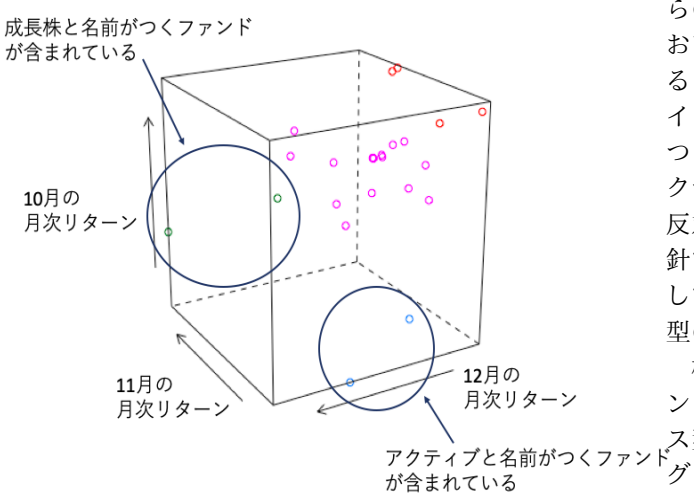


図 5: TOPIX をベンチマークとしている 25 のファンドを用いたクラスタリング結果

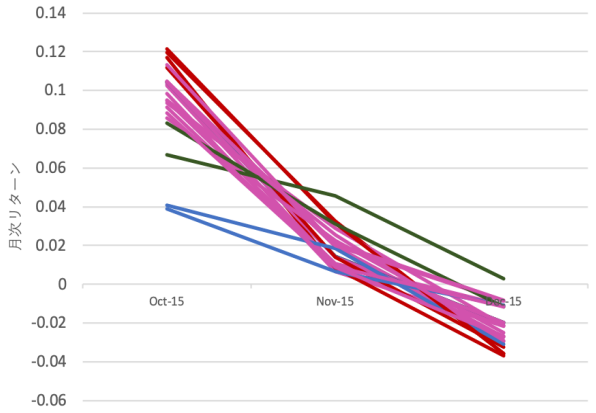


図 6: TOPIX をベンチマークとしている、25 のファンドの 10 月、11 月、12 月の月次リターンで、色はクラスタ番号を表す

で構成されているクラスタの中にはファンド名にインデックスが含まれるファンドが全て含まれていた。つまり最も大きな 17 つのファンドで構成されているこのクラスタは運用方針がインデックス型であるファンドのグループであると考えらえる。以上のことより、運用方針ごとにクラスタが形成されていると考えられる。

しかし、インデックス型であると考えられるクラスタの中にも、目論見書ではアクティブとしているファンドが含まれている。これらのファンドは二つの場合が考えらえる。一つ目はこの期間中だけインデックス型と同じリターンの値動きになってしまった場合であり、二つ目は本来はインデックス型の運用をしているにもかかわらず、アクティブ型と公表している場合である。これらのファンドは長期間観察することにより、どちらの場合なのか分類できると考えられる。どの期間においてもインデックス型の運用方針であると考えられるクラスタと同じクラスタに属している場合は、常にインデックス型の値動きと類似していると考えられる。つまり実際はインデックス型であるにもかかわらずアクティブ型と公表している可能性があると考えられる。反対に、他の期間においてはインデックス型の運用方針であると考えられるクラスタと異なるクラスタに属しているファンドは、運用方針に偽りなくアクティブ型の運用をしていると考えられる。

松井らの手法では同一の運用方針を取っているファンドのみを対象としていたが、本研究ではインデックス型とアクティブ型が混在するデータでクラスタリングを行い、運用方針が類似したファンドでクラスタが形成されたと考えられる。

6 まとめ

本研究では、実際の運用方針に基づくファンドのグループ分けを目的として、同一のベンチマークを対象として、月次リターンの類似度に基づいてファンドをクラスタリングする手法を提案した。

従来手法ではベンチマークが異なるファンドが混在したデータでクラスタリングを行っていたが、運用方針ではなく、ベンチマークごとにクラスタが形成された。本研究では、ベンチマークを統一してクラスタリングすることにより、運用方針でクラスタが形成されたことを確認した。

以上のことより、提案手法によって外部から観測できる月次リターンをもちいてクラスタリングした結果、アクティブ型と考えられるファンドが同じクラスタに含まれ、インデックス型と考えられるファンドも全て同じクラスタに含まれたので、運用方針ごとにクラスタが形成されたと考えられる。

参考文献

[1] 米田一樹, 松井藤五郎, 犬塚信博, 武藤敦子, 森山甲一:「投資割合を考慮した投資先類似度に基づく投資信託のクラスタリング」, 第 31 回人工知能学会全国大会 2D1-4(2017)

[2] 松井藤五郎, 米田一樹, 森山甲一, 武藤敦子, 犬塚信博:「リターン時系列に基づく投資信託の動的クラスタリング」, 第 33 回人工知能学会全国大会 IM4-J-13-04(2019)

t過程潜在変数モデルによるポートフォリオ生成

Portfolio Construction by t-Process Latent Variable Model

内山 祐介 [1,3*]　中川 慧 [2]
Yusuke Uchiyama[1,3] Kei Nakagawa[2]

[1] 株式会社 MAZIN
[1] MAZIN Inc.
[2] 野村アセットマネジメント株式会社
[2] Nomura Asset Management Co, Ltd.
[3] 筑波大学 システム情報系
[3] University of Tsukuba Faculty of Engineering, Information and Systems

Abstract: ガウス過程を潜在変数空間に拡張したものとして, ガウス過程潜在変数モデルが提案されており, 金融データの分析に使用した事例も報告されている. 一方で, 金融データは正規分布よりも裾の厚い確率分布に従うことから, ガウス過程潜在変数モデルによる分析では希少事象の生起確率を正確に推定できないことが懸念される. 本研究では, 裾の厚い確率分布に従うデータの分析に使用されるスチューデントの t-分布に基づいた, t-過程潜在変数モデルを提案する. このモデルを使用したポートフォリオ生成を行った結果, ガウス過程潜在変数モデルよりも良好な成績を収めた.

1　はじめに

機械学習で用いられる様々なアルゴリズムにおいて, データの生成モデルは特徴量が従う確率密度関数として与えられる. これらの学習モデルから確率変数としてスカラー量やベクトル量が抽出されるのに対して, 関数を抽出する生成モデルを考えることもできる. このような生成モデルは確率過程として知られており, 代表的なものはガウス過程である [1].

ガウス過程とは確率過程の一種であって, 任意に選択した有限個の確率変数の組の結合分布関数が正規分布で与えられる. 物理学や生物学の分野では, 時間変化する観測量のモデル化に確率過程が使用されてきたが, 入力変数は時間に限定されることなくどのような量であってもよい. 通常の正規分布と同様に, ガウス過程もまた平均と分散を与えれば確率密度関数の形状が一意に定まる. 通常の正規分布との違いは, ガウス過程の平均と分散は入力変数の関数として与えられる点である.

ガウス過程はまた, 重み係数が従う事前分布を分散が有限な確率密度関数で与えた際に, ベイジアンニューラルネットワークのノード数を無限大にとった極限として導出される. そのため, ガウス過程はニューラルネットワークと同等の表現力を有することが知られている [2]. 単層のニューラルネットワークに限らず, 深層ニューラルネットワークの極限としてもガウス過程が導出されることが報告されている [3]. 多くの学習モデルは入力に対する出力を点推定するのに対して, ガウス過程は平均関数による点推定に加えて分散関数による不確実性の推定も同時に行なえることが利点である.

近年, ガウス過程の金融分野への応用が積極的に進められている. 一例として, ポートフォリオ生成 [4] やコピュラに基づいた多変数間の動的相関の推定 [5], デリバティブの価格に関連するインプライド・ボラテリィリティの推定 [6] が挙げられる. これらの応用事例が報告されている一方で, ガウス過程は正規分布に基づいた学習モデルであるという点には注意する必要がある. 実際, 金融時系列は正規分布よりも裾が厚い確率分布に従うことが実証研究によって報告されている [7].

ガウス過程を裾が厚い確率分布に従うデータに拡張した学習モデルとして, t-過程が提案された [8]. このモデルから生成されたデータはスチューデントの t-分布に従うため, べき乗則に従って裾が広がる確率分布から生成されたデータのモデル化に適用することが可能である. t-過程は平均関数と共分散関数に加えて自由度と呼ばれるパラメータを与えると一意に決定される. 深層ニューラルネットワークの極限として得られるガウス過程を t-過程に拡張した, 深層 t-過程モデルが提案され金融時系列の解析に応用されている [9].

本研究では, t-過程を潜在空間に拡張した新たな学習モデルとして, t-過程潜在変数モデルを提案する. この

*連絡先： 株式会社 MAZIN
〒 111-0035 東京都台東区西浅草 3 丁目 29-14
E-mail: uchiyama@mazin.tech

モデルは裾が厚い確率分布に従うデータを潜在変数を使って表現することにより, 観測データのみを使用した従来の t-過程よりも広いクラスの問題に適用することが可能である. さらに, 金融の問題への応用として, t-過程潜在変数モデルによるポートフォリオ生成を実施した結果についても紹介する.

2 先行研究

2.1 ガウス過程

任意の有限個の $\{x_1, x_2, \cdots, x_n\}$ と確率過程 $f(\cdot)$ に対して, $\{f(x_1), f(x_2), \cdots, f(x_n)\}$ の同時確率密度関数が正規分布で与えられるとき, $f(\cdot)$ をガウス過程という. ガウス過程が従う確率密度関数は無限次元ベクトルが従う正規分布であるため, 平均値関数 $m(\cdot)$ と共分散関数 $K(\cdot,\cdot)$ を

$$m(x) = \mathbb{E}[f(x)], \tag{1}$$
$$K(x,x') = \mathbb{E}[(f(x)-m(x))(f(x')-m(x'))] \tag{2}$$

で定義すると, $X = [x_1, x_2, \cdots, x_n]^T$ が与えられたときの確率密度関数は $p(f|X) = \mathcal{N}(m(X), K(X,X))$ となる. 確率過程 f がガウス過程として抽出されることを明示する際には, $f \sim \mathcal{GP}(m(\cdot), K(\cdot,\cdot))$ と表記する. 簡単のため, ガウス過程の平均関数は 0 とされることが多い. 共分散関数はカーネル関数 $k(\cdot,\cdot)$ により,

$$K(x,x') = k(x,x') \tag{3}$$

で与えられる. カーネル関数は非負の対称関数であるため, $K(X,X)$ は正定値対称行列となる. カーネル関数としては動的基底関数

$$k_{\mathrm{RBF}}(x,x') = \alpha \exp(-l^{-2}||x-x'||^2) \tag{4}$$

が最もよく用いられている. 動的基底関数はハイパーパラメータとして α と l を持つ. これらは最尤推定等の最適化法によって決定される.

観測データの組 $\mathcal{D} = \{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$ に対して $X = [x_1, x_2, \cdots, x_n]^T$, $Y = [y_1, y_2, \cdots, y_n]^T$ とする. この (X,Y) を学習させたガウス過程から, 未知のデータ $X^* = [x_1^*, x_2^*, \cdots, x_m^*]^T$ に対する予測値 $Y^* = [y_1^*, y_2^*, \cdots, y_m^*]^T$ は, 平均関数 f^* と共分散関数 K^* がそれぞれ

$$f^* = m_X + K_{X^*,X}K_{X,X}^{-1}Y, \tag{5}$$
$$K^* = K_{X^*,X^*} - K_{X^*,X}K_{X,X}^{-1}K_{X,X^*} \tag{6}$$

で与えられる正規分布 $\mathcal{N}(f^*, K^*)$ から抽出される.

2.2 ガウス過程潜在変数モデル

与えられたデータのモデリングや分析を行う際に, より少ない次元でデータの特徴を捉えたいことがある. これは, 観測量 $y \in \mathbb{R}^D$ を $Q < D$ となるような潜在変数 $x \in \mathbb{R}^Q$ によって

$$y = f(x) + \varepsilon \tag{7}$$

と表現する際に, 写像 $f(\cdot)$ を推定する問題として定式化される. 特に, 観測ノイズ ε が正規分布に従い, かつ $f(\cdot)$ がガウス過程で与えられるときには, ガウス過程潜在変数モデルと呼ばれる [10]. N 個の観測データ $Y = [y_1, y_2, \cdots, y_N]^T$ に対して N 個の潜在変数 $X = [x_1, x_2, \cdots, x_N]^T$ を推定するときの尤度関数は

$$p(Y|X) = \frac{1}{(2\pi)^{ND/2}|K_{X,X}|^{D/2}} \exp\left(-\frac{1}{2}Y^T K_{X,X}^{-1} Y\right) \tag{8}$$

で与えられる. ここで, 平均関数は 0 とした. 尤度関数を潜在変数とカーネル関数のハイパーパラメータについて最適化することで, 観測データのガウス過程潜在変数モデルが得られる.

2.3 変分推論

式 (8) で与えられる尤度関数の最適化では過学習に陥りやすいことが知られている. そこで, $p(Y|X)$ の代わりに $p(X|Y) = p(Y|X)p(X)/p(Y)$ を最適化するベイズ学習が提案されている. 一般的に $p(Y) = \int p(Y|X)p(X)\mathrm{d}X$ を解析的に求めることは困難であるため, $p(X|Y)$ を X の適当な確率密度関数 $q(X)$ で近似することを考える. その際に KL 情報量

$$\mathrm{KL}[q|p] = \int \log \frac{q(X)}{p(X|Y)} q(X)\mathrm{d}X \tag{9}$$

を最小化する. 式 (9) はまた

$$\begin{aligned} \mathrm{KL}[q|p] &= -\int \log \frac{p(Y|X)p(X)}{q(X)} q(X)\mathrm{d}X \\ &+ \log p(Y) \end{aligned} \tag{10}$$

とも表現できるため, 右辺第 1 項についての最適化問題として定式化できることがわかる. ここで, $p(X)$ は潜在変数の事前分布である. この近似手法のことを変分推論という.

3 提案モデル

3.1 t-過程

ガウス過程では, 確率過程 $f(\cdot)$ が従う確率密度関数として正規分布を仮定していた. そのため, 金融時系列

のように正規分布よりも裾の厚い確率分布に従うデータに適用する際には, 正確な推定を行うことができなくなる. このようなデータに対してガウス過程を拡張したモデルが t-過程である [8]. t-過程は, 確率過程 $f(\cdot)$ が自由度 ν のスチューデントの t-分布

$$\begin{aligned}\mathcal{T}(m,K,\nu) &= \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{[(\nu-2)\pi]^{\frac{n}{2}}\Gamma\left(\frac{\nu}{2}\right)|K|^{\frac{1}{2}}} \\ &\times \left[1+\frac{1}{\nu-2}(y-m)^T K^{-1}(y-m)\right]^{-\frac{\nu+n}{2}}\end{aligned} \tag{11}$$

に従うモデルである. ここで, $m(\cdot)$ と $K(\cdot,\cdot)$ はそれぞれ平均関数と共分散関数であり, $\Gamma(\cdot)$ はガンマ関数である. 確率過程 $f(\cdot)$ が t-過程であるとき, $f\sim\mathcal{TP}(m(\cdot),K(\cdot,\cdot);\nu)$ と表記する. ガウス過程同様に, t-過程の平均関数も一般性を損なうことなく 0 として扱われることが多い.

t-過程の予測分布もまた t-分布となる. そのときの自由度, 平均関数および共分散関数は以下のように更新される [9].

$$\nu^* = \nu + n, \tag{12}$$

$$m^* = m + K_{X^*,X}K_{X,X}^{-1}Y \tag{13}$$

$$K^* = \frac{\nu-\beta-2}{\nu-n-2}\left[K_{X^*,X^*} - K_{X^*,X}K_{X,X}^{-1}K_{X,X^*}\right] \tag{14}$$

ガウス過程とは異なり, t-過程ではデータ数の影響が自由度と共分散関数の更新式に反映されていることが確認できる.

3.2 t-過程潜在変数モデル

ここでは, ガウス過程同様に t-過程を潜在変数モデルに拡張する. 観測データ $y\in\mathbb{R}^D$ と潜在変数 $x\in\mathbb{R}^Q$ とが t-過程 $f\sim\mathcal{TP}(m(\cdot),K(\cdot,\cdot);\nu)$ によって $y=f(x)$ と関係づけられることを仮定し, N 個の観測データを並べたものを $Y\in\mathbb{R}^{D\times N}$, N 個の潜在変数を並べたものを $X\in\mathbb{R}^{Q\times N}$ としたときに

$$\begin{aligned}p(Y|X) &= \frac{\Gamma\left(\frac{\nu+D}{2}\right)}{[(\nu-2)\pi]^{\frac{D}{2}}\Gamma\left(\frac{\nu}{2}\right)|K_{X,X}|^{\frac{1}{2}}} \\ &\times \left[1+\frac{1}{\nu-2}(Y-m_X)^T K_{X,X}^{-1}(Y-m_X)\right]^{-\frac{\nu+D}{2}}\end{aligned} \tag{15}$$

で与えられるモデルを, t-過程潜在変数モデルと定義する. スチューデントの t 分布は $\nu\to\infty$ の極限で正規分布に収束するため, t-過程潜在変数モデルはガウス過程潜在変数モデルを内包することがわかる. また, ガウス過程潜在変数モデル同様に, t-過程潜在変数モデルにおいても一般性を損なうことなく平均関数を $m(\cdot)=0$ とすることができる.

潜在変数とハイパーパラメータは観測データに対する最尤推定によって求められる. 式 (16) に対応する対数尤度関数は

$$\begin{aligned}\log p(Y|X) &= \log\Gamma\left(\frac{\nu+D}{2}\right) - \frac{D}{2}\log[(\nu-2)\pi] \\ &- \log\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\log|K_{X,X}| \\ &- \frac{\nu+D}{2}\log\left(1+\frac{1}{\nu-2}\Delta\right), \qquad (16)\\ \Delta &= (Y-m_X)^T K_{X,X}^{-1}(Y-m_X) \qquad (17)\end{aligned}$$

で与えられる. ガウス過程同様に, t-過程潜在変数モデルに対しても式 (9) で導入された変分推論を実施することが可能である. その際に, 潜在変数の事前分布と近似分布の具体形は分析者が決定することができる.

4 実証分析

ここでは, 代表的な世界各国の株式指数を用いて t-過程潜在変数モデルおよび比較手法としてガウス過程潜在変数モデルによるポートフォリオ構築を行う. 構築するポートフォリオは先行研究同様に最小分散ポートフォリオとする. すなわち, ポートフォリオの重み係数を w とすると, 式 (8) および式 (16) の共分散行列を表す $K_{X,X}$ を用いて

$$w^* = \arg\min_w w^T K_{X,X} w \tag{18}$$

を求める. t-過程潜在変数モデルによる最小分散ポートフォリオを Port_t, ガウス過程潜在変数モデルによる最小分散ポートフォリオを Port_G とする.

データセットとしては 16 か国の株価指数を用いた. それぞれ SP500 指数 (US), SP トロント 60 指数 (Canada), FTSE100 指数 (UK), CAC40 指数 (France), DAX 指数 (Germany), IBEX35 指数 (Spain), FTSE MIB 指数 (Italy), AEX 指数 (Netherlands), OMX30 指数 (Sweden), SMI 指数 (Switzerland), 日経平均株価 (Japan), 香港ハンセン指数 (HongKong), ASX200 指数 (Australia), 韓国 200 種株価指数 (Korea), OBX 株価指数 (Norway), MSCI Singapore 指数 (Singapore) である. データ期間は 1998/1/30 から 2019/1/31 までの約 20 年分, 253 サンプル (月次) である. データは Bloomberg 端末より取得した. 各指数の統計量は表 1 の通りである. 共分散行列 $K_{X,X}$ の推定のための条件を次の通り設定した.

推定期間: 10 年分 (120 サンプル)

テスト期間: 11 年分 (132 サンプル)

カーネル関数: Exp カーネル

表 1: 使用した株式指数の月次リターンの統計量

指数	US	Canada	UK	France	Germany	Spain	Italy	Netherlands
年率リターン	6.00%	5.41%	2.39%	4.08%	6.87%	3.20%	1.35%	2.96%
年率リスク	14.93%	14.92%	13.62%	18.12%	21.13%	20.66%	21.71%	19.13%
R/R	0.40	0.36	0.18	0.23	0.33	0.15	0.06	0.15
歪度	-0.66	-0.92	-0.55	-0.38	-0.50	-0.17	0.03	-0.74
尖度	5.23	7.36	4.53	4.52	6.12	4.96	4.80	5.88

指数	Sweden	Switzerland	Japan	HongKong	Australia	Korea	Norway	Singapore
年率リターン	6.32%	2.80%	3.35%	7.27%	4.70%	12.98%	10.72%	5.05%
年率リスク	19.51%	14.68%	19.24%	23.46%	12.40%	28.80%	21.49%	21.71%
R/R	0.32	0.19	0.17	0.31	0.38	0.45	0.50	0.23
歪度	-0.19	-0.73	-0.54	0.28	-0.69	1.39	-0.93	-0.26
尖度	5.29	6.11	4.75	5.78	4.54	11.63	6.84	6.81

表 2: ガウス過程潜在変数モデル, および t-過程潜在変数モデルに基づいて構成された最小分散ポートフォリオのパフォーマンス比較.

	Port_G	Port_t	差
前半 (2008/1∼2013/6)			
年率リターン	-4.89%	-2.63%	2.25%
年率リスク	19.57%	18.33%	-1.24%
年率 R/R	-0.25	-0.14	0.11
後半 (2013/7∼2019/1)			
年率リターン	6.08%	6.30%	0.22%
年率リスク	11.16%	10.56%	-0.60%
年率 R/R	0.54	0.60	0.05
全期間 (2008/1∼2019/1)			
年率リターン	0.64%	1.87%	1.23%
年率リスク	15.92%	14.93%	-0.99%
年率 R/R	0.04	0.12	0.09

潜在変数: 1 個

ここで, Exp カーネルは

$$k_{\mathrm{Exp}}(x, x') = \alpha \exp(-l^{-2}||x - x'||) \qquad (19)$$

で定義されるカーネル関数である. 以上の条件のもと, Port_t と Port_G の年率リターン, リスク, および R/R を比較する. 表 2 は, テスト期間を前半, 後半, および全期間に分けた際のパフォーマンス比較結果である. 前半はリーマンショックを含む下落局面であり, 後半は世界的な金融緩和に伴い株式市場の上昇局面であった. 表 2 から, 前半の下落局面ならびに後半の上昇局面のいずれにおいても, リターン, リスク, および R/R のすべての指標について, t-過程潜在変数モデルはガウス過程潜在変数モデルを上回っていることが確認できる.

5 まとめ

本研究では, t-過程を潜在空間に拡張した新たな学習モデルとして, t-過程潜在変数モデルを提案した. このモデルは裾が厚い確率分布に従うデータを潜在変数を使って表現することにより, 観測データのみを使用した従来の t-過程よりも広いクラスの問題に適用することが可能である. さらに, 提案手法の有効性を評価するための実証分析として, ガウス過程潜在変数モデルをベンチマークとした最小分散ポートフォリオ構築を世界各国の株価指数をもとに行った. 分析の結果, 提案手法である t-過程潜在変数モデルは相場の上昇局面, 下落局面に関わらずにガウス過程潜在変数モデルを上回った.

参考文献

[1] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Advanced lectures on machine learning*, pp. 63–71. Springer, 2004.

[2] Radford M Neal. Priors for infinite networks. In *Bayesian Learning for Neural Networks*, pp. 29–53. Springer, 1996.

[3] Jaehoon Lee, Jascha Sohl-dickstein, Jeffrey Pennington, Roman Novak, Sam Schoenholz, and Yasaman Bahri. Deep neural networks as gaussian processes. In *International Conference on Learning Representations*, 2018.

[4] Rajbir S Nirwan and Nils Bertschinger. Applications of gaussian process latent variable models in finance. In *Advances in Intelligent Systems and Computing*, pp. 1209–1221, 2019.

[5] David Lopez-Paz, José Miguel Henández-Lanato, and Zoubin Ghahramani. Gaussian process vine copulas for multivariate dependence. In *International Conference on Machine Learning*, pp. II10–II18, 2013.

[6] Yue Wu, José Miguel Henández-Lanato, and Zoubin Ghahramani. Gaussian process volatility model. In *Advances in Neural Information Processing Systems*, 2014.

[7] Benoit B Mandelbrot. The variation of certain speculative prices. In *Fractals and scaling in finance*, pp. 371–418. Springer, 1997.

[8] Amar Shah, Andrew Wilson, and Zoubin Ghahramani. Student-t processes as alternatives to gaussian processes. In *Artificial Intelligence and Statistics*, pp. 877–885, 2014.

[9] 中川慧, 角屋貴則, 内山祐介. 金融時系列のための深層 t 過程回帰モデル. 第 21 回 人工知能学会 金融情報学研究会 (SIG-FIN), 2018.

[10] Neil Lawrence. Gaussian process latent variable models for visualisation of high dimensional data. In *Advances in Neural Information Processing Systems*, 2004.

人工市場を用いた高頻度取引が金融市場流動性に与える影響の調査

Investigation of factors influencing market liquidity using artificial market

益田 裕司[1*]　水田 孝信[2]　八木 勲[3]
Yuji Masuda[1]　Takanobu Mizuta[2]　Isao Yagi[3]

[1] 神奈川工科大学大学院工学研究科情報工学専攻
[1] Course of Information and Computer Sciences, Graduate School of Kanagawa Institute of Technology
[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
[3] 神奈川工科大学情報学部
[3] Faculty of Information Technology, Kanagawa Institute of Technology

Abstract: 近年，金融市場では「流動性」に関心が集まっている．流動性は金融市場の盛況を表す尺度とされ、「取引のしやすさ」とも捉えられる．実証研究ではそれぞれの研究目的に沿った流動性指標が定義され，それらの有用性が議論されていた．この流動性に大きな影響を与えるといわれているのが高頻度取引（HFT）である．しかし，HFT が流動性指標にどのような影響を与えるのかは明確にされていない．そこで本研究では，人工市場を用いて，HFT を用いるエージェントが主要な流動性指標（Volume，Tightness，Resiliency，Depth）にどのような影響を与えるのかを指標の相関も考慮しながら調査した．その結果，HFT がいない場合と比べ，流動性指標は流動性が高まる方向へと変化していることがわかった．このことから HFT は市場流動性を高める効果がある可能性が示唆された．

1　まえがき

金融市場の盛況を表す目安とされる「流動性」に関心が集まっている．一般に流動性が高い市場とは，その時々で観察される『市場価格』に近い価格で，市場参加者が売りたい（あるいは買いたい）量を，速やかに売れる（あるいは買える）市場が想定されることが多い [1].

流動性に関する研究は，特に実証研究の分野で多数行われ，さまざまな知見が得られている [2][3][4][5]．例えば，Kyle[2] は，(1)”Tightness”，(2)”Depth”，(3)”Resiliency” を含む多くの概念を使用し流動性を評価することを提案した．Amihud[3] は，1 日の株式収益率の絶対値を取引金額で割った値を非流動性の尺度として提案し (ILLIQ)，資産の期待収益が流動性によって増加することを示した．

しかし，何をもって流動性とみなすかは実証研究の調査目的ごとに異なることが多い．例えば，市場価格のボラティリティの大きさや，市場参加者の売買が市場価格に大きな影響を及ぼさないことを流動性と結びつけることも多い．その結果、流動性を計測するために使用される指標は研究間で異なることが多く、ある研究で得られた流動性の結果が他の研究で得られた結果と一致しているかどうかを分析することは困難である．

一方で，この流動性に影響を与える要因の一つとして，HFT（High Frequency Trading，高頻度取引）があげられる．HFT が高速で取引を繰り返すことにより，取引量を増加させるため，金融市場を活発にさせる効果があるといわれている一方，市場に与える影響が大きいため暴騰暴落の頻度が高くなったともいわれている．さらに HFT が流動性にどのような影響を与えているか，体系的にほとんど議論されていない．

このように実証研究では対処困難なものに対応する手法の 1 つに，人工市場を用いる手法がある．人工市場は，社会シミュレーションの 1 つであり，計算機上に仮想的に構築されたマルチエージェントシステムの

*連絡先：神奈川工科大学大学院工学研究科情報工学専攻
神奈川県厚木市下荻野 1030
E-mail: s1885003@cco.kanagawa-it.ac.jp

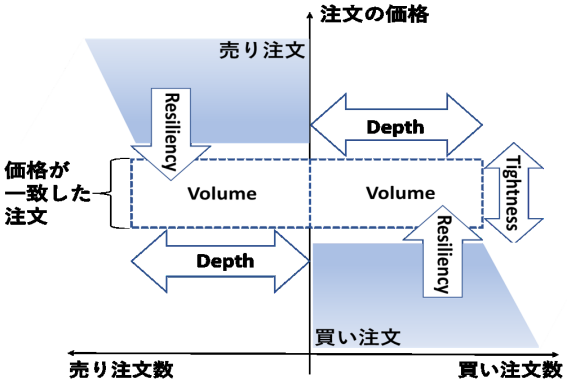


図 1: 流動性の概念整理 [9]

金融市場のことを指す [6][7][8]．人工市場におけるエージェントは仮想的な投資家であり，現実の投資家の特性がモデルとして組み込まれている．そして，エージェントらに金融資産の取引をさせることで市場がどのように振舞うかを確認することができる．また，市場側に規制や制約 (例えば，空売り規制のような市場の安定性と効率性を確保するための制限) をモデル化して組み込むことで，エージェントの振る舞いや市場にどのような影響が現れるかを検証することもできる．

人工市場を用いた HFT に関する研究は、いくつか存在する．しかしながら、HFT が流動性そのものにどのような影響を与えるかを調べた研究は見当たらない．

そこで本研究では，HFT を導入した人工市場のパラメータを変更して流動性指標間の関連を調査する．すなわち，HFT を導入した人工市場において，市場価格のティックサイズなどの市場内のパラメータを変更することで，HFT の存在が主要な流動性指標（Volume，Tightness，Resiliency，Depth）にどのような影響を与えるのか，HFT 非導入時の流動性指標の特性との比較などを行い，調査する．

本論文の構成は以下のとおりである．2 章では市場流動性とは何かについて説明し，流動性の実証研究についても説明する．3 章では本研究で用いた人工市場モデルについて説明する．4 章では本研究で行う実験の詳細や得られた結果について説明する．最後に 5 章では，本研究のまとめと今後の課題について述べる．

2 流動性

2.1 流動性の定義

流動性を計測する実証研究では，図 1 のように 4 つの評価軸（Volume，Tightness，Resiliency，Depth）がよく使用されている [9]．

2.1.1 Volume

第 1 の評価軸として，市場の取引量（Volume）があげられる．市場の取引量が多ければ，頻繁な取引や短期間での大口の取引がより容易になると考えられる．そのため，Volume が大きければ流動性が高いといえる．

Volume の指標としては，出来高，売買回転率をみる方法と，取引がない時間間隔，ゼロリターン率をみる方法の 2 種類が存在するが，本研究では出来高を Volume とする．

2.1.2 Tightness

第 2 の評価軸は，買い手と売り手の提示価格の差（Tightness）があげられる．提示価格の差が狭ければ，市場参加者の意図する価格から離れず取引が行える．そのため Tightness が小さければ流動性が高いといえる．

Tightness の指標としては，買い手の 1 番高い注文の価格である最良買い気配（ベスト・ビッド）と売り手の 1 番安い注文である最良売り気配（ベスト・アスク）の乖離幅として定義されるビッド・アスク・スプレッドがよく用いられる．本研究においてもビッド・アスク・スプレッドを Tightness とする．

2.1.3 Resiliency

第 3 の評価軸は，市場の復元力（Resiliency）があげられる．取引が成立し，上下した市場価格が元の価格に戻る早さのことを指す．価格に大きな騰落が起きた場合でも，迅速に騰落前の価格へ戻ると，取引を迅速かつ円滑に行える．そのため Resiliency は小さければ，流動性が高いといえる．

Resiliency の指標としては，値幅・出来高比率，ベスト・ビッドの枚数回復速度，ILLIQ といったものが存在する．

値幅・出来高比率は，日中の市場価格の最高値と最小値の幅を 1 日の出来高で除した指標で，その日の取引での平均的な価格変化を示している．本研究では値幅・出来高比率を Resiliency とする．

2.1.4 Depth

第 4 の評価軸は，市場の厚み（Depth）があげられる．ベスト・ビッド，ベスト・アスクに近い価格で出されている注文量であり，現在の市場価格に影響を与えることなく取引ができる数量を示す．ベスト・ビッド，ベスト・アスクに近い価格での注文量が多い状況を市場が厚い状況と呼ぶ．市場に厚みがあれば，市場参加者が意図した価格と市場価格との差が小さくなる．また市場に残っている注文が多く，取引成立した際の市

場価格が振れにくくなる．そのため，Depth が大きければ，流動性が高いといえる．

現実世界での Depth は，最良気配から 5，8，10 ティック離れた Depth 情報が公開され，流動性指標として利用されるようになった．本研究ではベスト・ビッドから 50 離れた価格までの注文量を Depth とする．

3 人工市場モデル

3.1 市場構成

本研究では，Mizuta.et.al[10] の人工市場モデルを基に，人工市場モデルの構築をおこなった．また高頻度取引を行うエージェントは，草田ら [11] のポジションマーケットメイカーエージェントを基にモデル化した．通常の投資家をモデル化した Mizuta.et.al[10] のエージェントをノーマルエージェント，草田ら [11] のポジションマーケットメイカーエージェントを HFT エージェントと呼ぶ．

本モデルは，1 つの資産のみを取引対象とし，ノーマルエージェントは n 体おり，ノーマルエージェント $j = 1$ から，$j = 2, 3, 4, \cdots$ と順番に注文を出す．最後のノーマルエージェント $j = n$ が注文を出すと，次の時刻にはまたはじめのノーマルエージェント $j = 1$ が注文を出していく．時刻 t はノーマルエージェント 1 体が注文を出すたびに，1 だけ増える．つまり，注文をしただけで取引が成立しない場合も 1 ステップ進む．HFT エージェントは各ノーマルエージェントが注文を出す前に，買いと売りの注文を行う．HFT エージェントの注文では時刻 t は進まない．このモデルでの価格決定メカニズムは買い手と売り手が価格を提示し，両者の提示価格が合致するとその価格で取引が成立する，連続ダブルオークション方式（ザラバ方式）とした．ティックサイズを ΔP とし，注文価格を求めるときに ΔP より小さい端数は買い注文の場合は切り捨て，売り注文の場合は切り上げる．買い注文価格より安い売り注文，または売り注文より高い買い注文が注文板に既に存在していれば，取引が即時成立する．取引が成立しなかった場合は注文を残す．本研究では，取引が即時成立する注文を成行注文，市場に残る注文を指値注文と呼ぶ．市場に残した指値注文が注文有効期間 t_c だけ経過しても取引が成立せず残っていた場合，注文板から取り除く．注文数は常に 1 とした．各エージェントが持つキャッシュ量は無限とし，資産を何単位でも買うことができる．また空売りも可能とした．

3.2 ノーマルエージェントの注文プロセス

ノーマルエージェントは以下の手順に従い，注文価格，買いと売りの判断を行う．ノーマルエージェント j が時刻 t のときに予想する価格の変化率（予想リターン）$r^t_{e_j}$ は式（1）で求める．

$$r^t_{e_j} = \frac{1}{w^t_{1,j} + w^t_{2,j} + u_j}(w^t_{1,j} \log \frac{P_f}{P^{t-n}} + w^t_{2,j} r^t_{h_j} + u_j \epsilon^t_j) \tag{1}$$

ここで，$w^t_{1,j}$ は時刻 t におけるノーマルエージェント j の i 項目の重みであり，シミュレーション開始時に，それぞれ 0 から $w_{i,max}$ までの一様乱数で決める．この重みは，後で述べる学習プロセスにより変化する．u_j は時刻 t におけるノーマルエージェント j の 3 項目の重みであり，シミュレーション開始時に，それぞれ 0 から u_{max} までの一様乱数で決められ，シミュレーション中に変化することはない．P_f は時間で変化しない一定のファンダメンタル価格である．P^t は時刻 t における市場価格 (取引されなかった時刻では直近で取引された価格であり，時刻 $t = 0$ では $P^t = P_f$ とする) であり，ϵ^t_j は時刻 t, ノーマルエージェント j の乱数項で，平均 0, 標準偏差 σ_e の正規分布乱数である．$r^t_{h_j}$ は時刻 t に，ノーマルエージェント j が計測した過去リターンであり，$r^t_{h_j} = \log(P^{t-n}/P^{t-(n+\tau_j)})$ である．ここで τ_j は，シミュレーション開始時に 1 から τ_{max} までの一様乱数でノーマルエージェントごとに決める．

式 (1) の第 1 項目はファンダメンタル価値を参照し投資判断を行うファンダメンタル投資家の成分であり，ファンダメンタル価格と直前期の市場価格を比較し，市場価格が安ければプラス，高ければマイナスの予想リターンを表す．第 2 項目は過去の価格推移を参照し投資判断を行うテクニカル投資家の成分であり，過去のリターンがプラスならプラス，マイナスならマイナスの予想リターンを表す．第 3 項目はノイズの成分を表す．

予想リターン $r^t_{e_j}$ より予想価格 $P^t_{e_j}$ は式（2）で求める．

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

注文価格 $P^t_{o_j}$ は平均 $P^t_{e_j}$，標準偏差 P_σ の正規分布乱数で決める．ここで，P_σ は式（3）で求める．

$$P_\sigma = P^t_{e_j} \times \ Est \tag{3}$$

Est ($0 < Est \leq 1$) を便宜上,「ばらつき係数」と呼ぶ．買いと売りの判断は予想価格 $P^t_{e_j}$ と注文価格 $P^t_{o_j}$ の大小関係で決まる．

$$\begin{aligned} &P^t_{o_j} < P^t_{e_j} \text{ なら 1 単位の買い} \\ &P^t_{o_j} > P^t_{e_j} \text{ なら 1 単位の売り} \end{aligned} \tag{4}$$

3.3 学習プロセス

状況に応じて戦略を切り替えるという学習プロセスを Yagi ら [12] のモデルを参考にモデル化した．学習はノーマルエージェントごとに注文の直前におこなわれ，ファンダメンタル投資家のみの場合の予想リターンを $r^t_{e_{1,j}} = \log(P_f/P^{t-1})$，テクニカル投資家のみの予想リターンを $r^t_{e_{2,j}} = r^t_{h_j}$ とする．これら $r^t_{e_{i,j}}$ を学習期間のリターン $r^t_l = \log(P^t/P^{t-t_l})$ と比較し，式（5）のように $w^t_{i,j}$ を書き換える．

$$\begin{aligned} &\text{同符号なら，} w^t_{i,j} \leftarrow w^t_{i,j} + k_l |r^t_l| q^t_j (w_{i,max} - w^t_{i,j}) \\ &\text{異符号なら，} w^t_{i,j} \leftarrow w^t_{i,j} - k_l |r^t_l| q^t_j w^t_{i,j} \end{aligned} \tag{5}$$

ここで，k_l は定数，q^t_j は時刻 t，ノーマルエージェント j に与えられる 0 から 1 までの一様乱数である．式（5）では，価格変化の方向の予測が現実と一致した戦略の重みを引き上げ，外れている戦略の重みを引き下げるようにしている．また式（5）の学習プロセスの他に，確率 m で $w^t_{i,j}$ を 0 から $w_{i,max}$ までの一様乱数にて再設定を行う．

3.4 HFT エージェント

HFT エージェントはマーケットメイク戦略（買い注文と売り注文のそれぞれに約定しやすい価格で指値注文を行う）をとる投資家をモデル化したものである．草田らの人工市場では 2 資産の市場を対象としているが，本研究では 1 資産のみを対象とする．

HFT エージェントは 1 体のみ存在し，各ノーマルエージェントが注文を出す前に，毎期買いと売りの指値注文を 1 単位ずつ行う．また HFT エージェントが新たに注文を出す際には，約定せずに注文板に残った前回の注文をキャンセルし注文を出しなおす．そのため，HFT エージェントどおしの取引は起こらない．

HFT エージェントは以下の手順に従い，買い売りそれぞれの注文価格を決定する．時刻 t におけるベスト・ビッドを $P^{t,buy}$，ベスト・アスクを $P^{t,sell}$，HFT エージェント固有のスプレッドを θ_{pm} とする．そして HFT エージェントが抱えるポジションを s^t_{pm}，自身の資産ポジション考慮度を ω_{pm} とすると，買い注文価格 $P^{t,buy}_{o,pm}$ と売り注文価格 $P^{t,sell}_{o,pm}$ は式 (6)，式 (7)，式 (8) で求める．

$$P^{t,buy}_{o_pm} = P^t_{fv,pm} - \tfrac{1}{2}(P_f \times \theta_{pm}) \tag{6}$$

$$P^{t,sell}_{o_pm} = P^t_{fv,pm} + \tfrac{1}{2}(P_f \times \theta_{pm}) \tag{7}$$

$$P^t_{fv,pm} = (1 - \omega_{pm}(s^t_{pm})^3) \cdot \tfrac{1}{2}(P^{t,buy} + P^{t,sell}) \tag{8}$$

上記の式では HFT エージェントの注文が成行注文になってしまう恐れがあるため，それを防ぐため式 (9) の制約を設けた．

表 1: 基準となるパラメータ

パラメータ	値
n	1,000
$w_{1,max}$	1
$w_{2,max}$	10
u_{max}	1
τ_{max}	10,000
σ_e	0.06
Est	0.003
t_c	20,000
ΔP	1.0
P_f	10,000
m	0.01
t_g	1,000
θ_{pm}	0.0003
ω_{pm}	0.00000005

$$\begin{aligned} P^{t,buy}_{o_pm} &\geqq P^{t,sell} \\ P^{t,sell}_{o_pm} &\leqq P^{t,buy} \end{aligned} \tag{9}$$

式 (9) の成約時の HFT エージェントの注文価格は式 (10)，式 (11) のようになる．

$$P^{t,buy}_{o_pm} = \begin{cases} P^{t,sell} - \Delta P & (P^{t,buy}_{o_pm} \geqq P^{t,sell}) \\ P^{t,sell}_{o_pm} - (P_f \times \theta_p m) & (P^{t,sell}_{o_pm} \leqq P^{t,buy}) \end{cases} \tag{10}$$

$$P^{t,buy}_{o_pm} = \begin{cases} P^{t,buy}_{o_pm} + (P_f \times \theta_p m) & (P^{t,buy}_{o_pm} \geqq P^{t,sell}) \\ P^{t,buy} + \Delta P & (P^{t,sell}_{o_pm} \leqq P^{t,buy}) \end{cases} \tag{11}$$

4 シミュレーション結果

4.1 実験概要

3 章でモデル化した人工市場を用いて実験を行う．各パラメータ値を変更して流動性の評価軸（Volume，Tightness，Resiliency，Depth）の変動を検証する．

実験で用いるモデルの基準パラメータ値を表 1 に示す．このうち，変更するパラメータとその値を表 2 に示す．パラメータは 1 つ 1 つ変化させ，変更しないパラメータについては基準パラメータ値に固定している．各種パラメータ値でそれぞれ 10 試行行い，後述する Volume，tightness，Resiliency，Depth を算出し，以後それらの平均値を求める．また，ボラティリティについても算出する．シミュレーションは時刻 $t = t_{end} = 1,000,000$ までおこなった．

表 2: 変更するパラメータと設定

パラメータ	値				
ΔP	0.01	0.1	1.0	10	100
σ_e	0.02	0.04	0.06	0.08	0.1
$w_{1,max}$	1.0	3.0	5.0	8.0	10.0
$w_{2,max}$	1.0	3.0	5.0	8.0	10.0

表 3: Stylized Fact

尖度		0.0777
価格騰落率の 2 乗の 自己相関	lag1	0.120
	lag2	0.0556
	lag3	0.0503
	lag4	0.0455
	lag5	0.0434

4.2 モデルの妥当性

実験に入る前に本人工市場モデルの妥当性を検証した．

シミュレーションモデルは実証研究 [13]，[14] で得られている統計的性質（stylized fact）が満たされているかで判断される．今回は人工市場に用いられる代表的な stylized fact であるファット・テールとボラティリティ・クラスタリングを判断基準とした．

表 1 の基準パラメータ値での尖度と騰落率の 2 乗の自己相関の統計値を表 3 に示す．この表からわかるように，尖度が正を示しているため，ファット・テールを満たしている．また騰落率の 2 乗の自己相関は，ラグがある場合でも正の相関を保っており，ボラティリティ・クラスタリングを満たしていることがわかる．以上より，本人工市場は妥当性があることが示された．

4.3 実験結果

4.3.1 ティックサイズ ΔP 変更実験

ティックサイズ ΔP を，0.01，0.1，1.0，10.0，100.0 と変化させたときの Volume，Tightness，Resiliency，Depth の平均を表 4 の左側，HFT 非導入時の 4 指標の平均を表 4 の右側に示す．

表 4 より，HFT が導入されると流動性指標は以下のような傾向を示した．

Volume は HFT の有無によらず，どちらも減少傾向であるが，ΔP が 10，100 の時を除き，HFT 導入時では非導入時と比べ，値は大きくなっている．Tightness，Resiliency もどちらも増加傾向であるが，ΔP が 100 の時を除き，HFT 導入時の両指標は非導入時と比べ，値は小さくなっている．Depth においては傾向自体が逆転し，HFT 導入時には増加傾向から減少傾向へと転じ，値は大きくなった．ΔP が 100 の時には，すべての流動性指標が HFT 導入時と導入時で同じような結果になった．

4.3.2 ファンダメンタル成分の重みの最大値 $w_{1,max}$ 変更実験

ファンダメンタル成分の重みの最大値 $w_{1,max}$ を 1.0，3.0，5.0，8.0，10.0 と変化させたときの Volume，Tightness，Resiliency，Depth の平均を表 5 の左側，HFT 非導入時の 4 指標の平均を表 5 の右側に示す．

表 5 より，HFT が導入されると流動性指標は以下のような傾向を示した．

Volume は HFT の有無によらず，どちらも減少傾向だが，HFT 導入時では非導入時と比べ，値は大きくなっている．Tightness はどちらも減少傾向であるが，HFT 導入時では非導入時と比べ，値は小さくなっている．Resiliency においては傾向自体が逆転し，HFT 導入時には減少傾向から増加傾向へと転じ，値は小さくなっている．Depth はどちらも増加傾向だが，HFT 導入時では非導入時と比べ，値は大きくなっている．

4.3.3 テクニカル成分の重みの最大値 $w_{2,max}$ 変更実験

テクニカル成分の重みの最大値 $w_{2,max}$ を 1.0，3.0，5.0，8.0，10.0 と変化させたときの Volume，Tightness，Resiliency，Depth の平均を表 6 の左側，HFT 非導入時の 4 指標の平均を表 6 の右側に示す．

表 6 より，HFT が導入されると流動性指標は以下のような傾向を示した．

Volume は HFT の有無によらず，どちらも減少傾向だが，HFT 導入時では非導入時と比べ，値は大きくなっている．Tightness はどちらも減少傾向であるが，HFT 導入時では非導入時と比べ，値は小さくなっている．Resiliency においては傾向自体が逆転し，HFT 導入時には減少傾向から増加傾向へと転じ，値は小さくなっている．Depth はどちらも増加傾向だが，HFT 導入時では非導入時と比べ，値は大きくなっている．

4.3.4 ノイズ成分の予想リターン計算 ϵ_j^t に用いる定数 σ_ϵ 変更実験

ノイズ成分の予想リターン計算 ϵ_j^t に用いる定数 σ_ϵ を 0.02，0.04，0.06，0.08，0.10 と変化させたときの Volume，Tightness，Resiliency，Depth の平均を表 7 の左側，HFT 非導入時の 4 指標の平均を表 7 の右側に示す．

表 4: HFT 導入時 (左) と非導入時 (右) における ΔP 変更実験

	HFT 導入時				HFT 非導入時			
ΔP	Volume	Tightness	Resiliency	Depth	Volume	Tightness	Resiliency	Depth
0.01	281,449	2.938	0.00157	3,961	265,884	12.015	0.0104	2,362
0.1	284,172	3.042	0.00164	3,885	267,309	12.174	0.0106	2,326
1.0	277,638	4.012	0.00230	3,804	266,209	12.406	0.0108	2,363
10.0	255,839	13.826	0.00776	3,053	266,892	17.579	0.0118	2,431
100.0	249,194	100.001	0.01933	3,183	249,419	100.001	0.0193	3,189

表 5: HFT 導入時 (左) と非導入時 (右) における $w_{1,max}$ 変更実験

	HFT 導入時				HFT 非導入時			
$w_{1,max}$	Volume	Tightness	Resiliency	Depth	Volume	Tightness	Resiliency	Depth
1.0	277,638	4.012	0.00230	3,804	266,209	12.406	0.0108	2,363
3.0	257,887	3.947	0.00233	4,171	247,682	12.014	0.0094	2,807
5.0	239,392	3.886	0.00238	4,510	230,026	11.772	0.0084	3,283
8.0	216,706	3.804	0.00248	4,926	207,453	11.171	0.0074	3,896
10.0	205,187	3.763	0.00255	5,139	196,318	10.675	0.0070	4,201

表 7 より，HFT が導入されると流動性指標は以下のような傾向を示した．

Volume は HFT の有無によらず，どちらも増加傾向だが，HFT 導入時のでは非導入時と比べ，値は大きくなっている．Tightness はどちらも増加傾向であるが，HFT 導入時では非導入時と比べ，値は小さくなっている．Resiliency においては傾向自体が逆転し，HFT 導入時には増加傾向から減少傾向へと転じ，値は小さくなっている．Depth はどちらも減少傾向だが，HFT 導入時では非導入時と比べ，値は大きくなっている．

4.3.5 ばらつき係数 Est 変更実験

エージェントごとの注文価格のばらつきを決めるばらつき係数 Est を 0.003，0.005，0.01，0.02，0.03 と変化させたときの Volume，Tightness，Resiliency，Depth の平均を表 8 の左側，HFT 非導入時の 4 指標の平均を表 8 の右側に示す．

表 8 より，HFT が導入されると流動性指標は以下のような傾向を示した．

Volume と Depth は HFT の有無によらず，どちらも減少傾向であるが，HFT 導入時の両指標は非導入時と比べ，値は大きくなっている．Tightness においては傾向自体が逆転し，HFT 導入時には増加傾向から減少傾向へと転じ，値は小さくなっている．Resiliency はどちらも増加傾向だが，HFT 導入時では非導入時と比べ，値は小さくなっている．

4.4 HFT 導入時と非導入時における流動性指標の変化

表 4 から表 8 より，HFT エージェントが導入されている市場では，導入されていない市場と比べ，Volume，Depth の値は増加し，Tightness，Resiliency の値は減少しており，すべての流動性指標が流動性を高める傾向にある．つまり，HFT を行う投資家が存在することにより，流動性は高くなる可能性が示唆されている．

Volume が高くなる理由として，HFT エージェントの注文価格設定方法があげられる．HFT エージェントは毎期の注文の際にベスト・ビッド，ベスト・アスクを参照して買いと売りの注文を行うため，それまでノーマルエージェントの注文で形成されていたベスト・ビッド，ベスト・アスクが HFT エージェントの注文に置き換わる．その結果，その後のノーマルエージェントの注文が成行注文になりやすく，出来高は増加傾向となる．

Tightness が減少する理由は，HFT エージェントがベスト･ビッド，ベスト･アスクを更新するような価格で買い注文と売り注文を出しているからである．HFT エージェントが出す買い注文と売り注文の幅は $P_f \cdot \theta_{pm} = 10000 \cdot 0.0003 = 3$ となるため，ビッド・アスク・スプレッドは大幅に小さくなる．

Resiliency が減少する理由は，値幅・出来高比率の分母は出来高であるが，HFT エージェントが約定しやすい価格で注文を出すため，出来高は増加する．HFT エージェントが出す注文が約定する場合，ボラティリティが低くなり価格の変動が抑えられたと考えられる．そのため，分子の 1 日の値幅は小さくなる．分母は大

表 6: HFT 導入時 (左) と非導入時 (右) における $w_{2,max}$ 変更実験

	HFT 導入時				HFT 非導入時			
$w_{2,max}$	Volume	Tightness	Resiliency	Depth	Volume	Tightness	Resiliency	Depth
1.0	406,155	4.497	0.00222	1,441	372,930	78.297	0.0518	223
3.0	366,397	4.330	0.00211	2,172	348,262	31.672	0.0267	645
5.0	333,144	4.208	0.00212	2,783	320,592	21.473	0.0180	1,080
8.0	296,810	4.080	0.00224	3,453	285,224	14.869	0.0123	1,840
10.0	277,638	4.012	0.00230	3,804	266,209	12.406	0.0108	2,363

表 7: HFT 導入時 (左) と非導入時 (右) における σ_ϵ 変更実験

	HFT 導入時				HFT 非導入時			
σ_ϵ	Volume	Tightness	Resiliency	Depth	Volume	Tightness	Resiliency	Depth
0.02	152,011	3.471	0.00335	6,116	151,109	5.308	0.0061	5,655
0.04	229,627	3.837	0.00254	4,693	222,479	8.975	0.0081	3,644
0.06	277,638	4.012	0.00230	3,804	266,209	12.406	0.0108	2,363
0.08	308,131	4.120	0.00218	3,239	296,523	15.928	0.0137	1,648
0.1	329,797	4.193	0.00216	2,832	318,035	18.821	0.0169	1,268

きくなり，分子は小さくなるため，値幅・出来高比率が小さくなる．

Depth が増加する理由は，HFT エージェントが導入されている市場では HFT エージェントの指値注文が約定しやすく，Depth の計測範囲内のノーマルエージェントの指値注文が消えないため，Depth は増加傾向にある．

ただし，ティックサイズ ΔP が 100 の時では上記の限りではない．このパラメータ値では，HFT 導入時と非導入時で流動性指標に変化がなかった．これは ΔP が 100 以外の場合，HFT 導入時のビッド・アスク・スプレッドはＨＦＴ非導入時よりも小さいが，ティックサイズが 100 になることでビッド・アスク・スプレッドは強制的に 100 に広げられる．ノーマルエージェントの注文幅が広くなったことで，HFT エージェントの指値にノーマルエージェントの注文が当たらなくなり約定しないため，HFT エージェントがいてもいなくても同じ状況にあると考えられる．

5 まとめ

流動性の代表的な 4 つの評価軸（Volume，Tightness，Resiliency，Depth）を計測するための代表的な指標に注目して，高頻度取引がそれら指標にどのような影響与えるのかをティックサイズなどの人工市場内のパラメータを変化させることで調査をした．その結果，どのパラメータを変更した場合においても，HFT が導入されていない市場の 4 指標の値と比べ，4 指標すべてが流動性が向上したといえる方向に変化した．このことから HFT を行う投資家の存在により，流動性が高くなる可能性が示唆された．

留意事項

本論文はスパークス・アセット・マネジメント株式会社の公式見解を表すものではありません．すべては個人的見解です．

参考文献

[1] T. Kurosaki, Y. Kumano, K. Okabe, and T. Nagano. Liquidity in JGB markets: An evaluation from transaction data. Bank of Japan Working Paper Series 15-E-2, Bank of Japan, May 2015.

[2] Albert S. Kyle. Continuous auctions and insider trading. *Econometrica*, Vol. 53, pp. 1315–1336, 1985.

[3] Yakov Amihud. Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, pp. 31–56, 2002.

[4] J. Muranaga. Dynamics of market liquidity of Japanese stocks: An analysis of tick-by-tick data

表 8: HFT 導入時 (左) と非導入時 (右) における *Est* 変更実験

	HFT 導入時				HFT 非導入時			
Est	Volume	Tightness	Resiliency	Depth	Volume	Tightness	Resiliency	Depth
0.003	277,638	4.012	0.00230	3,804	266,209	12.406	0.0108	2,363
0.005	220,858	3.984	0.00290	3,344	206,866	12.937	0.0125	2,405
0.01	146,416	3.942	0.00445	2,247	131,180	13.605	0.0191	1,821
0.02	84,989	3.898	0.00758	1,292	73,462	14.208	0.0191	1,088
0.03	59,677	3.879	0.01074	896	51,052	14.598	0.0191	754

of the tokyo stock exchange. In Bank for International Settlements, editor, *Market Liquidity: Research Findings and Selected Policy Implications*, Vol. 11, pp. 1–25. Bank for International Settlements, 1999.

[5] K. H. Chung, K. A. Kim, and P. Kitsabunnarat. Liquidity and quote clustering in a market with multiple tick sizes. *Journal of Financial Research*, Vol. 28, No. 2, pp. 177–195, 2005.

[6] S.-H. Chen, C.-L. Chang, and Y.-R. Du. Agent-based economic models and econometrics. *Knowledge Engineering Review*, Vol. 27, No. 2, pp. 187–219, 2012.

[7] C. Chiarella, G. Iori, and J. Perelló. The Impact of heterogeneous trading rules on the limit order flows. *Journal of Economic Dynamics and Control*, Vol. 33, No. 3, pp. 525–537, 2009.

[8] T. Mizuta, K. Izumi, I. Yagi, and S. Yoshimura. Regulations' effectiveness for market turbulence by large erroneous orders using multi agent simulation. In *Computational Intelligence for Financial Engineering Economics (CIFEr), 2014 IEEE Symposium Series on Computational Intelligence on*, pp. 138–143, March 2014.

[9] 土川顕, 西崎健司, 八木智之. 国際市場の流動性に関連する諸指標. Bank of Japan Review 15-E-2, 2013.

[10] T. Mizuta, K. Izumi, I. Yagi, and S. Yoshimura. Investigation of price variation limits, short selling regulation, and uptick rules and their optimal design by artificial market simulations. *Electronics and Communications in Japan*, Vol. 98, No. 7, pp. 13–21, 2015.

[11] 草田裕紀, 水田孝信, 早川聡, 和泉潔. 保有資産を考慮したマーケットメイク戦略が取引所間競争に与える影響:人工市場アプローチによる分析, 2015.

[12] I. Yagi, A. Nozaki, and T. Mizuta. Investigation of the rule for investment diversification at the time of a market crash using an artificial market simulation. *Evolutionary and Institutional Economics Review*, Vol. 14, No. 2, pp. 451–465, 2017.

[13] R. Cont. Empirical properties of asset returns: Stylized facts and statistical issues. *Quantitative Finance*, Vol. 1, pp. 223–236, 2001.

[14] M. Sewell. Characterization of financial time series, 2006.

強化学習を用いた期待効用ベースヘッジ手法

Expected Utility Based Hedge with Reinforcement Learning

上田 翼 [1*]
Tsubasa Ueda[1]

[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

Abstract: Selling options is a popular investment strategy, which regularly receives a premium and, on the other hand, takes variance risk, especially negative fat-tail risk. Therefore, it is important for risk-averse investors to mitigate these types of risks by constructing hedge position in consideration of transaction costs. Main results of this research are as follows: (1) In a practical simulation, DDPG model with utility based reward suggests a better way of dynamic hedging compared to simple benchmarks. (2) As a real-world application to market data, this learned model successfully manages the short straddle portfolio of treasury futures options.

1 はじめに

投資戦略のパフォーマンスを測るうえでは、期待リターンとボラティリティを基準に考えるのが一般的である．ある種の仮定の下では、平均分散分析によってポートフォリオを最適化すると、投資家の期待効用が最大化されるためである．しかしながら、投資家の効用関数が二次以外、かつ投資収益が正規分布しないときには、平均分散分析が必ずしも効率的にならないことが知られている [1]．前者の二次効用関数は単純化された仮定であり、現実的な投資家の行動を考えると拡張の余地は大きい．後者のように収益分布が非正規となる典型例はオプションやボラティリティ・インデックス（VIX 等）のポジションであり、市場が急変すれば損益が非線形に拡大しうる．このような戦略を適切に評価するには、平均分散分析では不十分であり、収益分布の歪度等を考慮した期待効用で考える必要がある．

一方、近年の技術的な動向においては、オプション・リスクの最適ヘッジ量を推定する手段として、深層学習が注目されている．[2] では、Semi-RNN を用いて凸リスク測度に基づくヘッジが提案された。[3] は、非線形の取引コストを仮定したうえで、DQN によるヘッジを扱っている。他方、著者の印象では、モデルの目的関数は実務的な観点から先験的に定義されたものが多いほか、実証面でも数値実験にとどまり現実の市場や投資戦略を必ずしも反映していないケースが多い。

そこで本稿では、期待効用理論の観点から自然な形で報酬を設計したうえで深層強化学習モデルを構築し、実際の市場データと投資戦略を用いて学習の妥当性を検証する．

*E-mail: tsubasa.ud@gmail.com

2 シミュレーション

2.1 期待効用モデル

投資家は、次のような絶対的リスク回避度 γ の指数型効用関数をもつと仮定する．

$$u(x) = \frac{1}{\gamma}(1 - e^{-\gamma x}) \tag{1}$$

x の平均 μ の周りでテイラー展開を行ったうえで期待値を取ると、

$$\begin{aligned} E[u(x)] &= u(\mu) - \frac{1}{2}u''(\mu)E[(x-\mu)^2] \\ &\quad + \frac{1}{6}u'''(\mu)E[(x-\mu)^3] + \dots \\ &\approx \mu - \frac{\gamma}{2}\mu_2[x] + \frac{\gamma^2}{6}\mu_3[x] \end{aligned} \tag{2}$$

と書ける。ここで、(2) は $\mu = 0$ 周りの近似であり、$u(\mu) \approx u(0) + u'(0)\mu = \mu$ を用いた．なお、μ_n は n 次中心モーメントである。ある T 時点の富を w_T とし、(1) を参照点 w_0 に依存する効用関数と捉えると、期待効用 $E[u(x)]$ の最大化は

$$\max\Big[E[w_T - w_0] - \frac{\gamma}{2}\mu_2[w_T - w_0] + \frac{\gamma^2}{6}\mu_3[w_T - w_0]\Big] \tag{3}$$

と表せる．$w_T - w_0 = \sum_{t=1}^{T} \Delta w_t$ と分解し、$Cov(\Delta w_t, \Delta w_s) = 0\,(t \neq s)$ を仮定すると、モーメントの加法性から、(3) は

$$\max \sum_{t=1}^{T}\Big[E[\Delta w_t] - \frac{\gamma}{2}\mu_2[\Delta w_t] + \frac{\gamma^2}{6}\mu_3[\Delta w_t]\Big] \tag{4}$$

と変形できる．[4] の議論を参考にすると、$\mu = E[\Delta w_t]$ が相対的に小さいならば $\mu_2 = E[(\Delta w_t)^2] - \mu^2$, $\mu_3 = E[(\Delta w_t)^3] - 3\mu\mu_2 - \mu^3$ の μ^2, μ^3 を無視できて、(4) は次のような報酬を設定した強化学習の問題として近似できる．

$$r_t = \mu - \frac{\gamma}{2}(\Delta w_t)^2 + \frac{\gamma^2}{6}\{(\Delta w_t)^3 - 3\mu(\Delta w_t)^2\} \quad (5)$$

ただし、μ は事前には明らかでないので後に置き換える．

2.2 環境とモデル

$t = 0$ 時点でオプションのポジションを構築し、離散時間 t ごとに原資産でヘッジする戦略を想定する．原資産価格が Heston モデル [5] に従うとして、状態空間と行動空間を次のように定義する．

$$\begin{aligned} s_t \in \mathcal{S} &:= \mathcal{R}_{++}^2 \times \mathcal{R}_+ \times \mathcal{R} \\ &= \{(S, \nu, \tau, \delta) \mid S, \tau > 0, \delta \geq 0, -1 \leq \delta \leq 1\} \\ a_t = a_t^\pi(s_t) &\in \mathcal{A} := \mathcal{R} \end{aligned}$$

S, ν, τ, δ はそれぞれ、原資産価格、原資産価格のインプライドボラティリティ、オプションの残存期間、ヘッジ比率を表す．a は定常な決定的方策 π に従って選択されたヘッジ資産の売買量であり、$\delta_{t+1} = \delta_t + a_t$ という関係が成立する．a_t から生じる取引コストを c_t、ヘッジなしオプション戦略の期待リターンを μ_{un} と表し、ヘッジ資産の期待収益を中立とみなせば、$\mu \approx \mu_{un} - E[c_t]$ と近似できるから、期待効用ベースの報酬 (5) を

$$\begin{aligned} r_t = &- c_t - \frac{\gamma}{2}(\Delta w_t)^2 \\ &+ \frac{\gamma^2}{6}\{(\Delta w_t)^3 - 3(\mu_{un} - c_t)(\Delta w_t)^2\} \end{aligned} \quad (6)$$

と置き換える．なお、$Cov(c_t, (\Delta w_t)^2) = 0$ を仮定し、μ_{un} を $a_t = 0$ $(0 \leq t \leq T)$ とおいて事前に推定しておくものとする．以上の環境は、連続状態行動空間のマルコフ決定過程 $M(\pi)$ となる．

目的関数として次の期待報酬和を設定し、DDPG(Deep Deterministic Policy Gradient) 法 [6] を適用する．

$$E\Big[\sum_{t=1}^{T} r_t \mid M(\pi)\Big] \quad (7)$$

DDPG の Actor および Critic は深層ニューラルネットワークで構成し、方策勾配法によりパラメータを更新する．

2.3 データセットと戦略

投資対象として米国債券先物オプション、ヘッジ資産として米国債券先物を想定する．離散化した Heston モデルにより、原資産価格とオプション価格のサンプルデータを、学習用に 400,000 ステップ、テスト用に 400,000 ステップ生成した．次章の実証分析を念頭において、Heston モデルにはリスクプレミアムの要素を導入し [7]、パラメータは実際の市場に近いものを設定した．

1 エピソードは、$S_0 = 100 + \epsilon$ と微小なランダム性を加えて基準化したうえで、$t = 0$ 時点で行使価格 $K = 100$ のショートストラドル（同一行使価格でのプットとコールの売りポジション）を構築し、t ごとにヘッジ比率 δ_t を調整しつつ、満期 $T = 40$ まで保有して終了する．オプションの満期前行使は想定せず、満期時に ITM(In The Money) であれば原資産に転換されるが、同時にポジション解消に必要な売買を行うものとする．取引コスト c_t はヘッジ資産の売買量 a_t に比例すると仮定した．保有するオプションとヘッジ資産の価格変化を反映したポートフォリオ収益を Δw_t とする．$R = w_T - w_0$ とすれば、期待報酬和 (7) の最大化は、効用関数 (1) に基づく収益の期待効用 $E[u(R)]$ を近似的に最大化していると解釈できる．

2.4 学習結果

図 1 は、テストデータで 10,000 回シミュレーションした収益 R の分布である．比較対象として、ヘッジなし (UN) と Black モデルに基づくデルタ（オプションの価格変化／原資産の価格変化）を毎時点で中立化する方法 (BD) を用意した．UN では、当然ながらボラティリティが大きく負のファットテールが目立っている．BD では、ボラティリティを大幅に抑制できているものの、取引コストを考慮せずヘッジ資産を頻繁に売買するため収益分布の中心が 0 に近づいている．強化学習モデルに基づくヘッジ（RL）では、収益分布の中心をある程度プラスに維持したままボラティリティとファットテールを適度に抑制できている．

同じことを統計量で確認したのが、表 1 である．平均と分散は t を日次とみなして R/S_0 を年率換算し、平均効用 $\overline{u(R)}$ は (2) の近似式ベースで計算した．BD と比べて RL の収益は分散や歪度が多少劣後するものの、平均収益率が高いため平均効用で僅かに上回る結果となった．

図 (2)、(3) は $\delta_{t-1} = 0$ とした際の RL と BD のヘッジ資産売買量を比較したものである．$0 \leq \tau(t) \leq 0.1$, $-2 \leq (S_t - K) \leq 2$ の範囲における a_t（垂直軸）を表している．RL は BD より a_t の絶対値が小さく、

コストを考慮してヘッジ資産の調整幅を制限していることがわかる．また、BD の a_t は正負が対称的になっているが、RL の a_t は正側が相対的に大きくなっており、原資産の下落リスクより上昇リスクを重視していると解釈できる．

表 1: 収益率の統計量（シミュレーション）

	UN	BD	RL
平均（年換算、%）	0.94	0.65	0.87
分散（年換算、%）	2.80	0.30	0.57
歪度	-1.01	0.04	-0.25
平均効用	-0.167	0.047	0.048

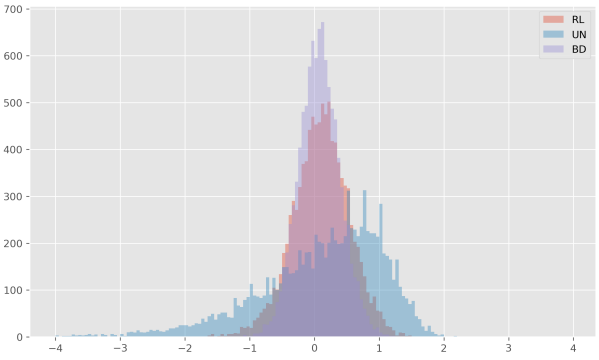


図 1: 収益分布（シミュレーション）

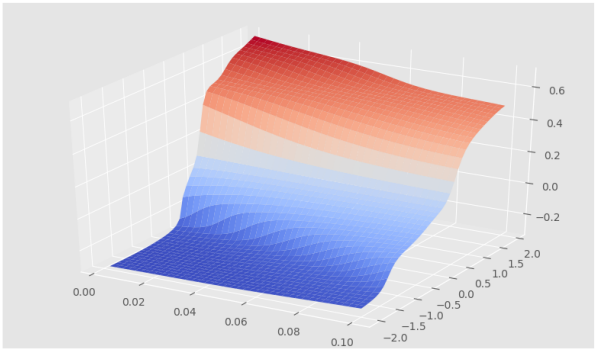


図 2: RL によるヘッジ調整量（シミュレーション）

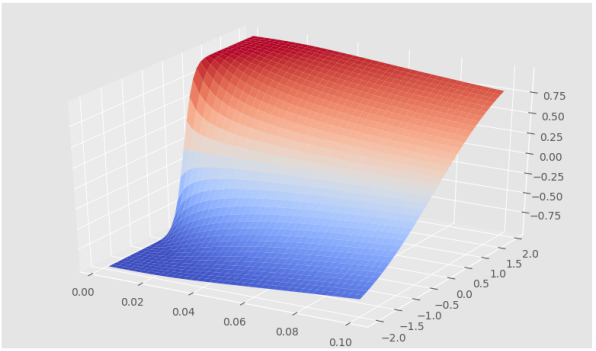


図 3: BD によるヘッジ調整量（シミュレーション）

3 実証分析

3.1 取引ルール

前章で学習した強化学習モデルを実際の市場データに適用しバックテストを行った．サンプル期間は 12 年 5 月～19 年 12 月とし、シミュレーションと同様に期近オプションを売り建てて日次でヘッジしつつ満期まで保有する戦略を繰り返した．なお、収益からはオプションやヘッジに用いる先物の売買コスト・手数料等を控除した．

3.2 バックテスト結果

図 4 は各ヘッジ手法を採用した際の NAV の推移である．UN と比べて RL のボラティリティやドローダウンは比較的抑制されており、長期的な累積リターンもプラスであることから、投資戦略として機能しているといえる．18 年後半からリターンがマイナス化しているが、実際に米金利のボラティリティが高止まりしていたため自然な結果である．

表 2 は、各ヘッジ手法の統計量である．前章のシミュレーション結果と同様に RL が最も高い平均効用を獲得した．なお、BD では平均収益率がほぼ 0 まで低下しており、現実的なヘッジ戦略を考えるうえで取引コストの重要性を示唆している．

表 2: 収益率の統計量（バックテスト）

	UN	BD	RL
平均（年率、%）	1.00	0.05	0.46
分散（年率、%）	2.79	0.72	0.93
歪度	-1.25	0.07	-0.44
平均効用	-0.106	-0.007	0.020

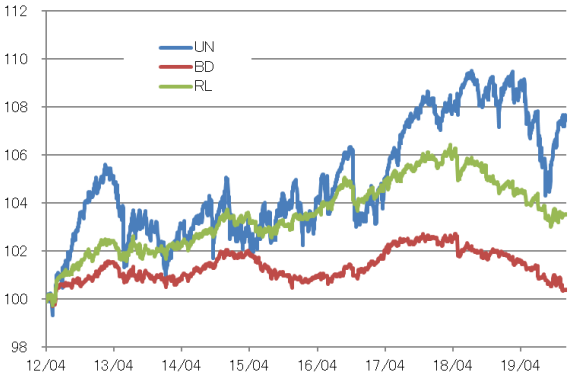


図 4: NAV の推移（バックテスト）

参考文献

[1] 安達智彦, 池田昌幸: 『長期投資の理論と実践』東京大学出版会 (2019)

[2] Buehler, Hans, et al. "Deep hedging." *Quantitative Finance* 19.8 (2019): 1271-1291.

[3] Kolm, Petter N., and Gordon Ritter. "Dynamic replication and hedging: A reinforcement learning approach." *The Journal of Financial Data Science* 1.1 (2019): 159-171.

[4] Ritter, Gordon, Machine Learning for Trading (August 8, 2017). Available at SSRN: https://ssrn.com/abstract=3015609 or http://dx.doi.org/10.2139/ssrn.3015609

[5] Heston, Steven L. "A closed-form solution for options with stochastic volatility with applications to bond and currency options." *The review of financial studies* 6.2 (1993): 327-343.

[6] Lillicrap, Timothy P., et al. "Continuous control with deep reinforcement learning." *arXiv preprint arXiv:1509.02971* (2015).

[7] Bollerslev, Tim, Michael Gibson, and Hao Zhou. "Dynamic estimation of volatility risk premia and investor risk aversion from option-implied and realized volatilities." *Journal of econometrics* 160.1 (2011): 235-245.

強化学習による高頻度取引戦略の構築

The Construction of High Frequency Trading Strategy via Reinforcement Learning

小林弘幸[1*] 和泉潔[1] 松島裕康[1] 坂地泰紀[1] 島田尚[1]
Hiroyuki Kobayashi[1] Kiyoshi Izumi[1] Hiroyasu Matsushima[1] Hiroki Sakaji[1] Takashi Shimada[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: 近年ゲーム AI の領域の成功により強化学習の研究が活性化し, 金融市場においても将来価格の予測に留まらず取引戦略をシステマティックに開発するための枠組みとして強化学習のアプローチに注目が集まっている. 本研究では東京証券取引所に上場する銘柄のティックデータから高頻度取引戦略を構築するための強化学習アルゴリズムを提案する. ニューラルネットワークを関数近似器として利用した強化学習により日本株市場のトレーディング戦略の構築を行い, 戦略の収益性をバックテストで確認する.

1 はじめに

今日の自動取引システムの一般的な開発プロセスでは, 取引戦略を過去の一定期間において稼働させた場合に発生する仮想的な損益やシャープレシオ等の評価尺度を元に, 開発者が人手によって取引モデルのロジックや閾値などのパラメータのチューニングを行っている. あるいは投資尺度を目的関数としてモデルのパラメータを調整する探索ベースの最適化手法が採用されることもある.

このバックテストを繰り返すことで取引戦略を構築する方法は取引スパンの長短を問わず広く採用されているが, このアプローチには固有の限界も存在する.

第一に, 行動判断の良し悪しの判定及びそのフィードバックが不十分な点が指摘される. 人間のトレーダーであれば取引判断の可否は即座に判定され, その事実を学習することができるが, バックテストにおいては取引戦略に従って選択した注文行動の帰結は全期間の平均的なパフォーマンスによってのみ評価され, 各々の取引や注文送出結果を学習材料として利用することは想定されていない.

第二に, エージェント（＝取引アルゴリズム）が取り得た他の状態を考慮しないという点が挙げられる. エージェントの発注判断材料として, 他の市場参加者も観測可能な「現在及び過去の市場状況」（Market State）に加えて, 「ポジション残高」や「既存注文状況」等のエージェント固有の状況（Agent State）も利用されるのが通常であるが, ある時点において取り得る無数の可能性の中で, （それまでの取引履歴に依存して）偶然にも経由することとなったただ一つの Agent State と, 行った発注判断の結果のみがパフォーマンス評価対象となり, 他の全ての起こり得る Agent State と注文送出の可能性は考慮されることがない.

そこで本研究ではこの限界を克服するために強化学習アプローチを提案する. 強化学習は価値ベースの手法と方策ベースの手法に大別されるが, 本研究では価値ベースを採用する. [状態 s, 行動 a] のペアに対する価値関数 Q(s,a) という形で現在のエージェントが置かれている状態の下で選択し得る行動の価値を推定し, 価値が最大となる行動を選択する. この方法によりエージェントの行動の成否は毎回独立したデータセットとして直接学習対象とすることができる.

加えて, 強化学習においては価値関数の推定に必要なデータを生成するために, エージェントに最適でない行動も確率的に選択させる. エージェントの行動にこのランダム性を付与することで Agent State の系列に幅を持たせることができ, 取引戦略の構築という視点からは現実の限られたマーケットデータを余すことなく利用できる手法であると考えている.

2 関連研究

強化学習を日中の取引戦略の構築に用いる研究の歴史は長く,2001 年には既にマーケットメイク戦略を Sarsa や Actor-Critic の手法で学習する手法が提案されている (Nicholas(2001)[1]).

*連絡先：東京大学工学系研究科,
〒 113-8656 東京都文京区本郷 7-3-1
E-mail: d2019hkobayashi@socsim.org

実際の株式の取引データを用いた研究としては米国株の執行のための最適指値戦略を Q 学習ベースの手法で構築した Nevmyvaka(2006)[2] が挙げられる.

強化学習が再注目された 2010 年代中盤以降の研究としてはタイルコーディングベースの線形関数近似と適格度トレースを用いて欧州株式市場のマーケットメイク戦略の学習を行った Spooner(2018)[3] や高頻度の逆張り戦略を MLP を用いてオンライン学習で構築した Ganesh(2019)[4] が挙げられる.

本研究はこれらの先行研究の設定を参考に, Atari などのゲーム AI 領域を中心に近年急速に研究が進む深層強化学習の手法を取り入れたものとなっている.

3 提案方法

3.1 学習ターゲット

従来の金融市場研究においては将来価格の予測に焦点を当てる文献が主流であるが, イントラデイのトレーディングにおいては「指値注文を提示し約定を待つ」戦略が多く観測され, 予測精度に加えて, 予測値の利用側である注文メッセージの送信タイミングを動的に決定するための発注戦略の重要性も高い.

予測と戦略決定は開発プロセス上分離可能なので前者を前工程, 後者を後工程と捉え, 本研究では後工程である「発注戦略の学習」にフォーカスする. 前工程のアウトプットである予測値に関しては実際の将来リターンにノイズを載せた数値で代用する.

3.2 環境シミュレータ

強化学習においてはエージェントの行動の結果として生ずる「状態遷移」及び「獲得する報酬」を計算するための環境シミュレータが必要となる. この環境シミュレータとして, 実際の板情報と歩み値データを用いてティックデータが変わる毎に注文の約定を判定する. このシミュレーションはエージェント自身の注文が市場に与えるマーケットインパクトは考慮できない性質のものである.

3.3 強化学習アルゴリズム

本研究が想定する取引エージェントは連続値の状態空間と, 有限個で少数（7 個）の行動空間を取る. そこで Deep Q Network（DQN)[5] をベースとした強化学習アルゴリズムを採用する. [5] の研究は画像データが入力値のために CNN を用いているが, 本研究では少数（6 個）の「状態変数」を入力値とし, 全結合型ニューラルネットワークを関数近似器として採用する.

加えて, 現在 Atari のプレイスコアが最も高く SoTA とされる R2D2(2019)[6],Ape-X(2018)[7] の強化学習モデルが取り入れている初代 DQN に対する 3 つの拡張

- Double Q Nework [8]
- Prioritized Experience Replay [9]
- Dueling Network [10]

を適用したものを本研究の強化学習アルゴリズムとする.

3.4 状態

以下の 6 個の状態変数を定義する.[1]

1. 現在ポジション株数/最大保有株数
2. 既存買注文の反対気配との乖離率 (逆数)
3. 既存売注文の反対気配との乖離率 (逆数)
4. Bid-Ask スプレッド
5. 短期（5 秒）リターン予測値
6. 長期（60 秒）リターン予測値

最大保有株数は 100 万円相当の株数（最小取引単位の 100 株に満たない場合は 100 株）で, ショートポジションも取り得るものとしている.

3.5 行動

以下の 7 種類の行動を可能な行動集合とする.

1. 何もしない
2. 最良気配への新規買注文送出/既存買注文訂正
3. 最良気配への新規売注文送出/既存売注文訂正
4. 反対気配への新規買注文送出/既存買注文訂正
5. 反対気配への新規売注文送出/既存売注文訂正
6. 既存買注文の取消
7. 既存売注文の取消

取引エージェントの行動タイミングは各銘柄のティック更新時に, 前回行動から 1 秒以上経過している場合に発注判断を行うものとする (各発注判断を 1 ステップとする）. ある時点において市場に出すことのできる注文は買い, 売りのそれぞれに対して 1 注文のみとし, 最大ポジションを超える新規注文は行わない. 各時点で選ぶことのできない行動は選択肢から排除している.

[1] 状態 2, 状態 3 は既存注文が無い状態の数値を 0（乖離率は∞）として扱うために逆数を取ったものを状態変数としている.

3.6 報酬

1 ステップごとに発生する損益額に加え, ポジションが増大することによるペナルティ項とキャンセル注文送出によるペナルティ項を付加したものを報酬 r(t) として次の式を用いて算出する.

$r(t) = \Delta PnL(t+1) - \alpha(\Delta Pos(t+1))^2 - \beta Fo(t)$

$\Delta PnL(t+1)$: (累積) 損益額の変化幅（円)

$\Delta Pos(t+1)$: ポジション金額の変化幅 (円)

$$Fo(t) = \begin{cases} 1 & (cancel\ order\ was\ submitted\ at\ t) \\ 0 & (otherwise) \end{cases}$$

損益額のみを報酬とする場合, ポジションが必要以上に増加する傾向にあり [3], また不必要な注文の送信, 取消を繰り返してしまう. そこで過度なポジショニングと注文取消を抑制するために調整項を加えている.

3.7 ニューラルネット

本研究で関数近似器として用いるニューラルネットの入力層ユニット数は強化学習の「状態変数」の数の 6 , 出力層ユニット数は強化学習の「行動」の数の 7 となる.

Dueling Network[10] のアイディアを取り入れ,Q(s,a) を状態価値関数 Value(=V(s)) と行動による価値の増分 Advantage(=A(s,a)) に分解する.

$Q(s,a) = V(s) + A(s,a) - 1/n\Sigma^{n}_{a'=1}A(s,a')$

n : 行動の数 (= 7)

この分解で全てのサンプルデータを状態価値関数の学習に用いることができるようになり学習が安定する.

Dueling Network を実装するために 2 層の中間層を設定する (図 1). 第 1 中間層は全結合層とし, ユニット数は 20 とする. 第 2 中間層は V(s) の計算用と A(s,a) の計算用に分岐し, それぞれユニット数 10 に設定する. 活性化関数には $\tanh x = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$ を採用する.

3.8 パラメータの学習

行動価値関数 Q の値が最大となる行動のみを選択する評価用エージェント（Greedy 方策）1 体と, 確率 ε で取り得る行動からランダムに選択する探索用エージェント（ ε - Greedy 方策）を 5 体用意する. [2]

探索用エージェント i の行動履歴 $Experience_i[s,a,r,s']$ は Replay Buffer に蓄えられ, 1 日（1 エピソード）の終了後にバックプロパゲーションによりニューラルネット・パラメータ（ウェイト）の学習が行われる. [3] 次式で定義される TD 誤差 $\delta(\theta)$ の二乗和を減少させる方向に SGD でパラメータ θ の学習が行われる.

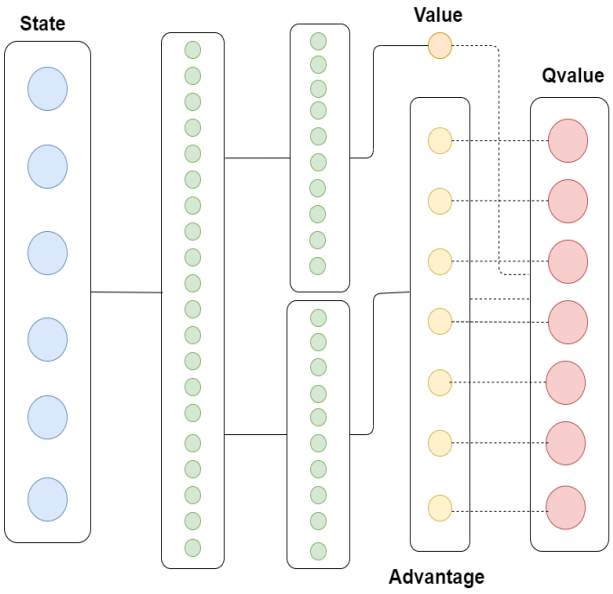


図 1: 本研究の関数近似器である Dueling Network.

$\delta(\theta) = (r + \gamma\, Q(s', a^*|\theta^*)) - Q(s,a|\theta)$

$a^* = argmax_{a'}\ Q(s', a'|\theta)$

r : 報酬

γ : 割引率 (0.999)

s, a : 状態, 行動

s', a' : 次ステップ（遷移先）の状態, 行動

θ : ニューラルネット・パラメータ

θ^* : 過去の時点のニューラルネット・パラメータ

$\theta \leftarrow \theta - \alpha\nabla_\theta\delta(\theta)^2/2 = \theta + \alpha\ \delta(\theta)\nabla_\theta Q(s,a|\theta)$

α : 学習率

この学習に用いるデータ (Experience) は Replay Buffer からランダムに取り出されるが, 本研究では Schaul(2016)[9] に倣い各 Experience の TD 誤差の絶対値をサプライズの大きさと捉え, 選択される確率を TD 誤差の絶対値に比例するように設定している.

学習の対象となる期間全体を 1 エポックとして日付順にシミュレーションと学習を行う. 各 Experience は次エポックの当該営業日に Replay Buffer から除去される. すなわち第 2 エポック以降は常に直近の 1 エポック分の記憶から学習データがサンプリングされることとなる. 平均を取ると 1 つの Experience が 8 回程度学習用にピックアップされるように SGD の繰り返し回数を調整している.

[2] 本研究はシングルエージェントシミュレーションで各エージェントの取引は他のエージェントが参照する板情報に影響を与えることは無い.

[3] 学習はエージェント毎ではなく, 銘柄毎に単独の学習器が全ての探索用エージェントの行動履歴を利用して行う.

4 実験と考察

4.1 データセットと取引対象

データは東京証券取引所のヒストリカルデータ（Flex Full) を利用し, 上下 10 本の板情報を再現する. 取引対象は東京証券取引所に上場する主要企業から 5 銘柄（武田製薬 (4502), ソニー (6758), トヨタ自動車 (7203), 三菱 UFJ フィナンシャルグループ (8306), ソフトバンクグループ (9984)）を選定する.

4.2 期間及び取引時間

学習期間は 2018 年 7〜12 月の 124 営業日を 1 エポックとして 20 エポック繰り返して学習を行う. 学習で得られたニューラルネット・パラメータを用いて 2019 年 1〜6 月のデータでバックテストを行い強化学習で獲得した戦略の収益性を確認する（テスト中と最終エポックはニューラルネット・パラメータの更新は行わない）.

東京証券取引所の取引時間の 9:00〜11:30（前場）,12:30〜15:00（後場）の中で各セッションの開始後 10 分と最後の 5 分は取引を行わず, 翌日にポジションを持越さないように 14:55 に成行でポジションを解消する. [4]

4.3 前提条件と評価方法

発注タイムラグに関してはゼロ（注文送出後次のティック時刻までに自注文が取引所に到達）と仮定している. 取引に要する取引料, アクセス料, 委託手数料, 品貸料等の諸経費は考慮しない.

評価は学習したエージェントの行動選択をグリーディ選択に固定して行う. 取引のシャープレシオは日次ベースで計算し年間平均営業日数 (246) の平方根を掛けて年率換算する.

予測精度がどの程度必要かを把握するために将来リターン予測値に載せるノイズの大きさを変えて評価を行う. ノイズは (-N, N) の一様分布とし, 区間長 N を調整することでノイズの大きさを調整する. 実際のノイズの大きさは短期, 長期のそれぞれの将来リターンの標準偏差を参考に 2N(bp)，5N(bp) としている.

4.4 学習過程

図 2 はソフトバンク株式（9984) に関して, ノイズ N=0（完全予見ベース）とした時の学習の様子を確認したものである. 横軸は経過エピソード数, 縦軸は各エピソードの損益と累計報酬を基に直近 1 周分（124 エピソード）のデータからシャープレシオを算出している.

[4]特別気配が出ている時間やセッションをまたぐ状態遷移は学習対象から外している.

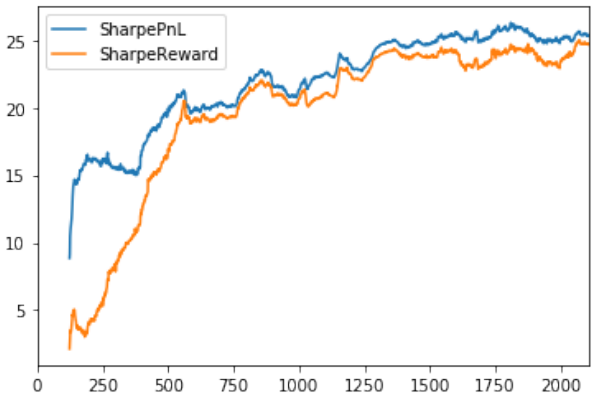


図 2: 学習過程（9984・完全予見ベース）

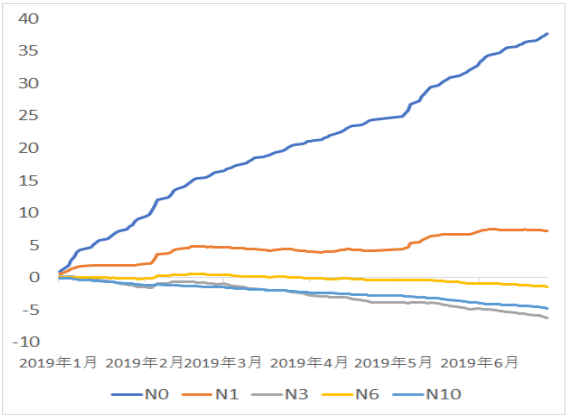


図 3: バックテスト期間中の累積損益

4.5 実験結果

学習期間で得られたニューラルネット・パラメータを用いてテスト期間でバックテストを行った結果を図 3, 表 1 に示す. 表 1 中の数値は 6 か月間 (117 営業日）の平均値となっている. また, 銘柄ごとにシャープレシオを算出した表を表 2 に示す.

4.6 考察

完全予見の設定ではエピソードを重ねるごとに学習が進んでおりアウトオブサンプルのテストセットにおいても収益性の高い行動選択が可能であった. 一方で, ノイズを付加していくにつれて戦略の構築に失敗する銘柄の割合が増加し, それに伴いポートフォリオ全体の収益性も低下する. ノイズ率とパフォーマンスの関係が逆転するケース (4502）も観測され, 学習が安定していない可能性が示唆される.

表 1: バックテスト結果.

ノイズ N	売買代金（百万円）	発注件数（件）	損益（円）	シャープレシオ
0	3,104	6,752	322,045	27.21
1	6,591	13,398	61,411	6.59
3	4,889	9,452	-53,556	-10.53
6	1,562	2,979	-11,965	-4.23
10	1,374	2,090	-40,500	-27.30

表 2: 銘柄ごとのシャープレシオ.

ノイズ	4502	6758	7203	8306	9984
0	21.18	20.44	19.37	29.03	17.60
1	-6.62	12.64	-29.34	-20.70	15.38
3	16.25	-8.68	-18.90	-70.88	7.97
6	-17.78	-18.38	-21.19	-12.00	9.81
10	-2.33	-17.38	-20.22	-40.16	-9.20

5 むすびと今後の展望

本研究では関数近似器にニューラルネットワークを用いた強化学習により日本株式の高頻度取引戦略の構築を試みた.

高精度の将来予測値を利用可能な状況化においては有効な取引戦略を獲得できることが確認できたが, 予測値のノイズレシオを増加させた時は収益性の低下が見られ, 現実的な予測精度の環境において適用するためには強化学習アルゴリズムの細部を調整する必要があると考えている.

また現在の 5 銘柄,6 エージェント,20 エポックの強化学習シミュレーションにおいて, ワークステーションのシングルスレッドで 1 日以上の学習時間を要すことから, 学習をスケールさせるために今後は並列計算に取り組むことを予定している.

参考文献

[1] Nicholas T. Chan and Christian R. Shelton. 2001. An Electronic Market-Maker. AI Memo 2001-005. MIT AI Lab.

[2] Yuriy Nevmyvaka et al.: Reinforcement Learning for Optimized Trade Execution. Proceedings of the 23 rd International Conference on Machine Learning, Pittsburgh, PA, (2006)

[3] Thomas Spooner et al.: Proc. of the 17th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2018), M. Dastani, G. Sukthankar, E. Andre, S. Koenig (eds.), July 10-15, 2018, Stockholm, Sweden.

[4] Ganesh, Prakhar & Rakheja, Puneet. (2018). Deep Reinforcement Learning in High Frequency Trading.

[5] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. Nature, 518(7540):529, 2015.

[6] Kapturowski, Steven, Georg Ostrovski, John Quan, Remi Munos, and Will Dabney. "Recurrent experience replay in distributed reinforcement learning." (2018).

[7] Horgan, Dan, John Quan, David Budden, Gabriel Barth-Maron, Matteo Hessel, Hado Van Hasselt, and David Silver. "Distributed prioritized experience replay." arXiv preprint arXiv:1803.00933 (2018).

[8] Van Hasselt, Hado, Arthur Guez, and David Silver. "Deep reinforcement learning with double q-learning." In Thirtieth AAAI conference on artificial intelligence. 2016.

[9] Schaul, Tom, John Quan, Ioannis Antonoglou, and David Silver. "Prioritized experience replay." arXiv preprint arXiv:1511.05952 (2015).

[10] Wang, Ziyu, et al. "Dueling network architectures for deep reinforcement learning." arXiv preprint arXiv:1511.06581 (2015).

機械学習による金融時系列モデルのパラメータ推定

Parameter estimation of financial time series models by machine learning

高石哲弥

Tetsuya Takaishi

広島経済大学
Hiroshima University of Economics

Abstract: 金融資産の実証分析において、ボラティリティは金融資産のリスクを表す重要な量である。ボラティリティの推定において、よく利用される手法は、資産価格の収益率時系列をモデル化し、ボラティリティを推定する方法である。このとき、モデルのパラメータは収益率時系列に合うように推定される。本研究では、金融時系列モデルの１つである GARCH モデルのパラメータ推定に機械学習の手法を用い、パラメータの推定が正しく行えることを示す。また、学習率の違いによる収束率についても述べる。

１．はじめに

金融における実証分析では、日々変動する金融資産のリスクを見積もり、そして将来のリスクを予測することは保有する金融資産を安全に管理するために必要不可欠である。リスクを見積もる代表的な指標としては価格変動の大きさを表すボラティリティがある。ボラティリティの変動を予測するための手法としては、価格時系列の性質を捉えたモデルを構築し、そのモデルから将来のボラティリティを予測する方法がある。金融資産価格の時系列の統計的性質は精力的に調べられており、その結果、多くの金融資産価格に共通に現れる性質があることが知られている。それらの性質は“Stylized facts”と呼ばれている[1]。Stylized facts として特に重要な性質は次である。

（１）収益率分布は裾野の厚いファットテイル分布である。
（２）収益率の自己相関は短期の相関である。
（３）絶対値収益率の自己相関は長期相関である。
（４）ボラティリティクラスタリングを示す。

実証分析では、多くの Stylized facts を再現でき、かつ予測精度の高いモデルが有用とされる。そのようなモデルの例として GARCH モデル[2-6]があり、実証分析でよく利用されている。その他に確率的ボラティリティ変動モデルもよく利用されるモデルである。

モデルを利用する場合、モデルのパラメータを時系列データにもっとも合うように推定する必要がある。そして、推定されたパラメータの下で将来のボラティリティ変動を予測する。

推定方法としては、最尤法やベイズ推定等が存在する。確率的ボラティリティ変動モデルのように、ベイズ推定が有効な推定法とされているモデルもある。ベイズ推定の実行には、マルコフ連鎖モンテカルロ法が良く利用される。そのため、有効なマルコフ連鎖モンテカルロ法の開発も重要な課題として研究がなされている。

本研究では、GARCH モデルのパラメータ推定を機械学習によって実行し、パラメータが正しく決定できるかどうかを調べる。

２．GARCH モデル

本研究では、GARCH モデルの中で最もシンプルな３つのパラメータを持つ GARCH(1,1)モデル[2]を用いる。価格時系列を r_t {t=1,...,N} とすると、GARCH(1,1)モデルは以下で記述される。

$$r_t = \sigma_t \varepsilon_t \qquad (1)$$

$$\sigma_t{}^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2 \qquad (2)$$

ここで、σ_t^2は分散、ε_tは正規乱数を表している。

このモデルにおいて、パラメータはα, β, ωの３つである。これらのパラメータを時系列データにもっとも合うように推定する。

３．ベイズ推定

ベイズの定理を利用することによって、時系列r_tが与えられたもとでパラメータ$\theta = (\alpha, \beta, \omega)$の従う確率分布$p(\theta \mid r_t)$が以下のように与えられる。

$$p(\theta \mid r_t) = \frac{L(r_t \mid \theta)\pi(\theta)}{f(r_t)} \quad (3)$$

ここで、$L(r_t \mid \theta)$は尤度関数、$\pi(\theta)$はパラメータの事前分布である。$f(r_t)$は規格化定数にあたるが、ベイズ推定をマルコフ連鎖モンテカルロ法で実行するときには、重要ではない。

GARCH モデルの$L(r_t \mid \theta)$は、以下で与えられる。

$$L(r_t \mid \theta) = \prod_{t=1}^{N} \frac{1}{(2\pi\sigma_t^2)^{1/2}} \exp\left(-\frac{r_t^2}{2\sigma_t^2}\right) \quad (4)$$

ベイズ推定において、パラメータの推定値は確率分布$p(\theta \mid r_t)$のもとでのパラメータの期待値$\langle \theta_i \rangle$として与えられる。ここで、$(\theta_1, \theta_2, \theta_3) = (\alpha, \beta, \omega)$と置いた。

$$\langle \theta_i \rangle = \int \theta_i p(\theta \mid r_t) d\theta \quad (5)$$

一般には、上記の積分は解析的には行えないので、マルコフ連鎖モンテカルロ法によって期待値$\langle \theta_i \rangle$を見積もる。機械学習による結果との比較のために、ベイズ推定よるパラメータ推定も実行する。マルコフ連鎖モンテカルロ法には様々な手法があり、GARCH モデルにもいくつかの手法が試されている[7,8]が、本研究では多変量スチューデント分布を提案分布に用いた Metropolis-Hastings 法を用いた[9-14]。

４．機械学習による推定

機械学習によるパラメータ推定が正しく行えるかどうかを見るために、まずパラメータを$\alpha, \beta, \omega = (0.3, 0.6, 0.05)$に設定した人工の時系列データを生成し、そのデータを利用して正しいパラメータ推定ができるかどうかを決定する。

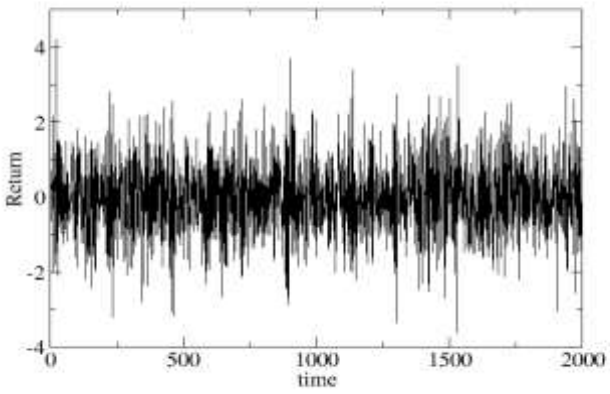


図１：人工時系列データ$\alpha, \beta, \omega = (0.3, 0.6, 0.05)$

図１は$\alpha, \beta, \omega = (0.3, 0.6, 0.05)$に設定し、（１）及び（２）式によって生成した時系列長 N=2000 のデータである。このデータから逆に生成に使ったパラメータを推定する。

モデルの対数尤度から損失関数を以下のように定義し、機械学習によって最適化を行う。

$$\mathrm{LOSS} = \sum_{t=1}^{N} \frac{1}{2}\left(\log(\sigma_t^2) + \frac{r_t^2}{\sigma_t^2} + \log(2\pi) \right) \quad (6)$$

最適化には Adam optimizer[15]を利用した。実行に際し利用した計算機環境は表１である。

表 1: 計算機環境

CPU	Intel Core i7-8700 3.20GHz
Memory	16GB
OS	Windows 10
TensorFlow[16]	Version 1.9.0
Python	Version 3.6.6

パラメータの初期値は$\alpha, \beta, \omega = (0.1, 0.8, 0.1)$に設定し、Adam optimizer の学習率(LR)は様々な値を用いてパラメータの収束を調べた。その結果、LR=0.0001～0.005 の値においては正しい値に収束することを確認した。そして、LR=0.005 を超

えるあたりから不安定となり収束しない結果となった。

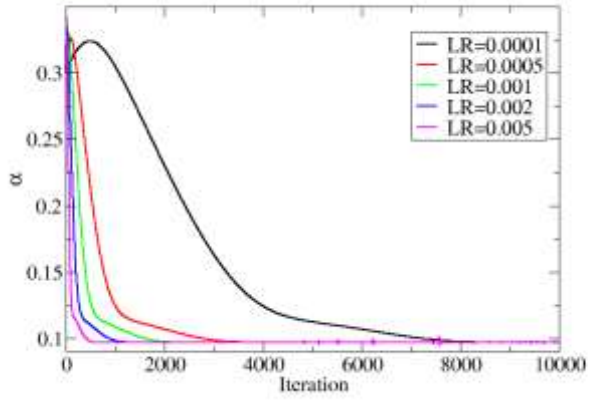


図 2：パラメータ α

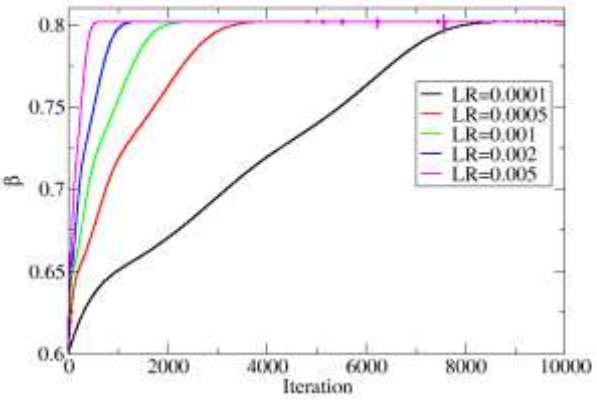


図 3：パラメータ β

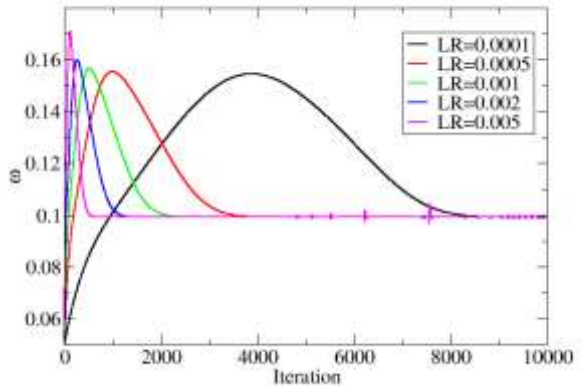


図 4：パラメータ ω

図 2～4 は様々な LR に対するパラメータ α, β, ω の収束の様子を表している。LR=0.005 まで、LR が大きいほど速く正しい値に収束している。

機械学習とベイズ推定による実行結果は、表 2 にまとめてある。機械学習によって求められた値と設定値は非常に似た結果となっており、機械学習によって正しい結果が求められている。

表 2 ：推定結果

	α	β	ω
設定値（真値）	0.1	0.80	0.1
機械学習	0.0967	0.802	0.0991
ベイズ推定	0.103	0.779	0.115

次に、学習率による収束の速さについて調べる。学習率ごとの収束の速さ（収束率）を以下のように測定した。まず、真値に収束した後の損失関数の値を最小値として、誤差を損失関数の最小値からの乖離として定義する。

Error=損失関数（iteration）―損失関数の最小値

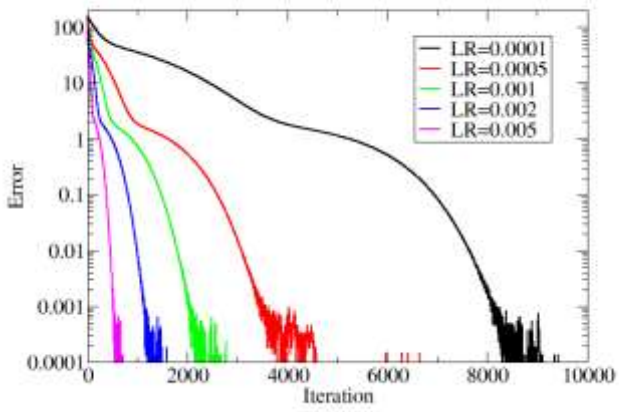


図 5：学習率の違いによる誤差減少の違い

図 5 はこのように定義した誤差がどのように減少していくかを表している。収束は LR が大きいほど早く、誤差の小さい領域では指数関数的に減少しているように見える。そこで、誤差の小さい領域で誤差を $a\exp(-\mu x)$ の関数でフィットする。ここで、a と μ はフィッテイングパラメータ、x は iteration である。フィッテイングによって求めた μ の値をプロットしたのが図 6 である。μ が大きいほど収束が速いことを意味するので、学習率が大きいほど収束が速くなることを示し

ている。但し、学習率が 0.005 を超えると不安定となり収束しない。

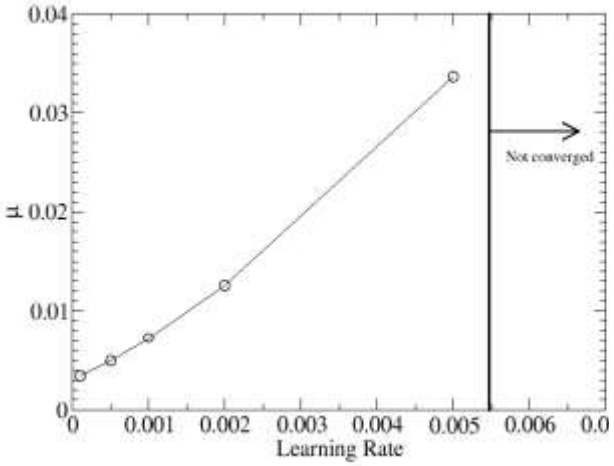


図 6：μ と学習率(LR)の関係

５．まとめ

本研究では、金融時系列モデルの１つである GARCH モデルのパラメータ推定を機械学習の手法（Adam optimizer）によって実行した。予め、パラメータ値の分かっている時系列を利用し、機械学習によるパラメータ推定を行った結果、正しい結果が得られることを確認した。また、Adam optimizer の学習率を変えて実行し、学習率が大きいほど正しい値への収束が速いことが分かった。但し、学習率がある値を超えると、収束が不安定となることも分かった。

参考文献

[１] R.Cont, Empirical properties of asset returns: stylized facts and statistical issues. Quantitative. Finance 1, 223–236, (2001)

[２] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307-327, (1986)

[３] R.F. Engle, “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom inflation,” Econometrica, vol.50, 1982, pp.987-1007.

[４] T. Bollerslev, R.Y. Chou and K.F. Kroner, “ARCH modeling in finance,” *Journal of Econometrics*, vol.52, 1992, pp.5–59.

[５] T. Takaishi, Rational GARCH model: An empirical test for stock returns,” *Physica A*, vol.473, 2017, pp.451-460.

[６] T.Takaishi, “Volatility estimations using a rational GARCH model”, Quantitative Finance and Economics, vol.2, pp.127-136.

[７] L. Bauwens, M. Lubrano, “Bayesian inference on GARCH models using the Gibbs sampler,” *Econometrics Journal*, vol.1, 1998, pp.23-46.

[８] T. Takaishi, “Bayesian Estimation of GARCH model by Hybrid Monte Carlo,” *Proceedings of the 9th Joint Conference on Information Sciences*, 2006, CIEF-214. Available: https://doi.org/10.2991/jcis.2006.159

[９] T. Takaishi, “An Adaptive Markov Chain Monte Carlo Method for GARCH Model,” *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering. Complex Sciences*, vol.5, 2009, pp.1424-1434. Available: https://doi.org/10.1007/978-3-642-02469-6_22

[１０] T. Takaishi, “Bayesian Estimation of GARCH Model with an Adaptive Proposal Density,” *New Advances in Intelligent Decision Technologies, Studies in Computational Intelligence*, vol.199, 2009, pp.635–643. Available: https://doi.org/10.1007/978-3-642-00909-9_61

[１１] T. Takaishi, “Bayesian Inference on QGARCH Model Using the Adaptive Construction Scheme,” *Proceedings of 8th IEEE/ACIS International Conference on Computer and Information Science*, 2009, pp.525-529. Available: https://doi.org/10.1109/ICIS.2009.173

[１２] T. Takaishi, “Bayesian inference with an adaptive proposal density for GARCH models,” *J. Phys.: Conf. Ser.*, vol.221, 2010, 012011.

[１３] T. Takaishi and T.T. Chen, “Bayesian Inference of the GARCH model with Rational Errors,” *International Proceedings of Economics Development and Research*, vol.29, 2012, pp.303–307. Available: http://www.ipedr.com/vol29/55-CEBMM2012-R00014.pdf

[１４] T. Takaishi, “Markov Chain Monte Carlo versus Importance Sampling in Bayesian Inference of the GARCH model,” *Procedia Computer Science*, vol.22, 2013, pp.1056 – 1064.

[１５] D.P. Kingma and J.L. Ba, “ADAM: A METHOD FOR STOCHASTIC OPTIMIZATION,” arXiv:1412.6980. Available: https://arxiv.org/pdf/1412.6980.pdf

[１６] TensorFlow https://www.tensorflow.org/

会社四季報を用いた理論株価の推定
~Kaggle を参考に~

Calculating Company's Stock Prices from The Japan Company Handbook

吉川　満 [1]

Mitsuru Kikkawa[1]

[1] 大和大学政治経済学部
[1] The Faculty of Political Science and Economics, Yamato University

Abstract: This paper calculates the theoretical stock prices in the fundamentals theory with "The Japan Company Handbook", which covers all listed companies. Three machine learning methods are used with referring the developed the analytical methods in Kaggle.

1. はじめに

株式の取引を行う際，全上場企業の企業概要や業績データ等が，コンパクトにまとめられている「会社四季報」[1]など情報ハンドブックは，投資家の間で，広く利用されている。会社四季報が発売された日は，「四季報相場」[2]と呼ばれるいつもとは異なる相場が形成されることがあるなど，影響力があることが知られている。ただし，これには，膨大な情報があり，個人投資家はどの項目が重要で，どのように株価に関係しているのか等まで理解するのは困難である。

近年，様々なデータが容易に利用できるようになってきたこともあり，機械学習[3]の様々な手法が開発されている。開発された手法を用いて，様々な Web サイトにて，株価や為替等の予測を試みているものも多く公開されている。また，Kaggle（カグル）と呼ばれるデータサイエンスに関するコミュニティが注目されている。特に Kaggle 中で，定期的に企業等がコンペ形式[4]で課題を提示し，参加者が予測精度を競い合っている。

本稿では，会社四季報にある構造化された情報を Kaggle 参加者の手法を参考にし，機械学習を活用することで，ファンダメンタルズ理論における理論株価[5]を予測する。

[1] 東洋経済新報社が年 4 回発行（新春号 12 月中旬，春号 3 月中旬，夏号 6 月中旬，秋号 9 月中旬）している。また，この会社四季報は，購入せずとも，証券会社が提供している取引ツールからも見ることができるものもある。

[2] 最近では既に会社四季報 ONLINE [1]において 2 週間程前から先出しの情報があり，事前に株価が値上がりしている銘柄も多いと指摘されている。

[3] 機械学習とは，人間が持つ学習にあたる仕組みを機械（特にコンピュータ）で実現する技術・手法の総称と言われている。一般的に，構造化されたデータを分析する際には，機械学習，非構造化されたデータを分析する際には，深層学習（ディープラーニング）が適していると言われている。

[4] コンペには，例えばメルカリによる販売者が投稿した情報を基に「適正な販売価格」を予測するもの[2]やリクルート社が運営する「ホットペッパー」と「Air レジ（飲食店向け POS レジアプリ）」のデータを利用して，レストランのお客さんの数を予測するものなどがある[3]。コンペ終了後，成績上位者が提案した分析モデルを見ることもできる。株価の予想に関するものでは，"Two Sigma: Using News to Predict Stock Movement" [4] というコンペがあり，Two Sigma Investments（米国ヘッジファンド）が主催のスポンサーをしている。このコンペでは，米国株式市場の日足の取引データと，Reuters が提供する経済ニュースを利用して，将来の株価（10 日先の株価）のリターンを予測するというコンペがある。

[5] 株価形成理論に関して，大きく分けて次の 3 つに分類することができる。① ファンダメンタルズ理論，② 効率的市場仮説，③ 需給理論

2. データと分析手順

2.1 会社四季報

会社四季報には，図 1 にあるように，会社基本情報，業績推移，財務詳細，株価詳細，株主情報等のデータと会社四季報編集部のコメントが全上場企業（3723 社，第 2 集）の（構造化された）情報が掲載されている。ただし，算出できない，非回答等に関しては，[-]となっている。

これらの情報から導出される理論株価とは，その企業に投資した場合，将来にわたり投資家へもたらす利潤を予想し，算出した計算上の株価水準のことを指している。現在，投資家の間で広く知られている財務・業績から算出できる理論株価の 2 つを付録にまとめている。これらの理論株価は，ある独自の計算式にこれらのデータを当てはめることで，導出している。また，このような理論株価は著名な投資家も投資家の多くが参考にして売買していると説明している[5]。

図 1　各企業の内容（CD-ROM 版，イメージ）

（出所　株式会社東洋経済新報社　ホームページ）

2.2 データ分析手順

Kaggle 等で活躍されている方の分析手順の概略[6]を次のようにまとめることができる。本稿でもこの手順を参考にして行う。

(1) 探索的データ分析（EDA，Exploratory Data Analysis）

データの可視化や相関分析等，簡単な分析を行い，データの特徴を明らかにする。

(2) ベースラインモデルの構築

最初に与えられた生データを基に有名なモデル，例えばランダムフォレストをそのまま適用するなど，容易に結果が得られるモデルの構築を行う。

①ファンダメンタルズ理論は，財務状況や業績をもとにして，企業の本質的な価値を分析する方法のことをいう。

[6] 例えば，[6]，[7]を参照されたい。

ただし，Kaggle では，モデルを提出し，公開されていない未知のデータでのテストの結果の精度を競い合うものであり，本稿の状況とは異なる。

(3) 交差検証（Cross Validation）
まず訓練データを分割し，その一部をまずベースラインモデルを用いて，予測精度の高い最適なモデルを得る。次に，訓練データのうち，利用していないデータで，その最適なモデルを用いて，テストを行い，モデルの妥当性の検証を行う。

(4) 特徴量エンジニアリング[8]
会社四季報のデータにおいても，産業・上場区分等の数値以外のデータを取り扱うため，数値に置き換える必要がある。
また，検証を行いながら，数値の対数化，変化率等にすることや2つ以上の種類の特徴量を組み合わせて，例えば，1株あたりに関するものを入れることで予測精度が上がるのであれば，新しい特徴量を作るなどを行う。ここでは，ドメイン知識と呼ばれる専門分野に関する知識が重要であると指摘されている。

(5) ハイパーパラメータ調整
モデルのパラメーターを調整し，予測精度を上げていく。

(6) アンサンブル
テストデータは未知であり，訓練データで最も予測精度が高いものが必ずしも，テストデータでも予測精度が高いとは限らない。そのため複数のモデルの予測を組み合わせて，全体の予測精度を上げていく。

3. 評価指標と特徴量エンジニアリング

本稿では，個人投資家が会社四季報を次のような使い方をしていると想定している。会社四季報が発売され，業績，状況等のファンダメンタルを精査し，今後どのような株価となるのかを予想する。そこから株価は，割安・割高なのかを判断する。
そこで，会社四季報のデータを機械学習，教師あり学習の手法[7]を用いて，理論株価を導出する。ここでは，訓練データを「2019年第2集(春号)」(2019年3月15日発売)とデータ作成時点の終値(2019年2月25日)，テストデータを「2019年第3集(夏号)」(2019年6月18日発売)における全上場企業の財務データ等（説明変数・特徴量）を用いて，データ作成時点の終値(2019年5月30日)の株価（終値，被説明変数）を推計する。ただしその日に取引がなく，終値がない場合，最後に取引が行われた日の株価を用いる。

3.1 評価指標

モデルの評価は，機械学習等の分野でよく使われている次の3つの指標を用いる(n:企業数)。

(1) 平均絶対誤差率(MAPE, Mean Absolute Percentage Error)

$$\mathrm{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{\left|\textit{理論株価}_i - \textit{実際の株価}_i\right|}{\textit{実際の株価}_i}$$

(2) 平均二乗誤差(MSE, Mean Square Error)

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\textit{理論株価}_i - \textit{実際の株価}_i)^2$$

(3) 決定係数（R^2）

実際の株価が理論株価に近い，予測精度が高い場合，このMAPE，MSEの値は小さく，R^2の値は1に近くなる。本稿では，MAPE，MSEの値が小さく，R^2の値が1に近いモデルを予測精度の高いモデルとする。

3.2 探索的データ分析

会社四季報のデータの主な特徴や相関分析の結果を次にまとめた。
(1) 会社四季報から得られる説明変数・特徴量は137あり，数が多いものの，会社数は3721(第2集(春号))と少ない。
(2) 最近，上場した会社など多くの会社において，欠損値が多い。
(3) 株価（終値，被説明変数）と説明変数・特徴量と相関分析の結果，相関係数(0.5～0.7)が高いものは，「1株益(連結)」，「1株当り純資産(連結)」，「1株配当(円)（連結)」であり，これらが高いほど株価も高い関係にあることが分かった。また，これらのデータは，既存の理論株価の導出に用いられている。

[7] 教師あり学習とは，答えを含めたデータセット（訓練データ）を用意し，それを用いて学習を行い，モデルを作成する。作成したモデルを元に主に分類や回帰を行う。

3.3 ベースラインモデル

ベースラインモデルとして，容易に推計できる重回帰分析を用いて行った。しかしデータに欠損値が多く，予測をすることができなかった。そのため欠損値の多い銘柄や説明変数・特徴量を削除した。その結果，MAPE は約 206.8 となり，予測精度は 2 倍，半分以上異なった。

3.4 特徴量エンジニアリング

説明変数・特徴量に関して，主に次のことを行った。

(1) 上場区分，業種，日経 225 採用銘柄などをカテゴリー化した。
(2) 株価と発行済株数に関しては，数値のばらつきが大きいため，対数化した。
(3) データには，マイナスの値があり，欠損値が多い。そこで，特徴量・データをカテゴリー化し，欠損値自体も 1 つのカテゴリーとして取り扱った。具体的には，データに順位をつけ，10%毎にカテゴリー化し，欠損値を 11 番目のカテゴリーとした。

4. 推定結果と考察

4.1 推定結果

ベースラインモデルとして，ランダムフォレスト，勾配ブースティング(2 種類)の計 3 つの手法にて，理論株価の推計を行なった。

特にこの勾配ブースティング（Gradient Boosting）は，予測手法として，近年注目されている決定木アルゴリズムに基づいた勾配ブースティングの機械学習フレームワークである（1）XGBoost と（2）LightGBM を用いる 。この勾配ブースティングは Kaggle コンテストで最も成功した最近の機械学習アルゴリズムの 1 つであると言われている [9]。また，実際 Kaggle 入賞者たちのアルゴリズムによく使われている。

3 つのモデルにおける結果は次のようになった。

(1) ランダムフォレスト

	train	test
MAPE	2.312	4.492
MSE	0.050	0.163
R^2	0.947	0.825

(2) XGBoost

	train	test
MAPE	5.014	5.615
MSE	0.216	0.249
R^2	0.767	0.731

(3) LightGBM

	train	test
MAPE	3.294	4.771
MSE	0.111	0.185
R^2	0.881	0.801

第 2 集(春号)と第 3 集(夏号)の終値に関する MAPE は約 13.3 であり，終値は，平均約 13.3%変動しており，上記 3 つのモデルにおいて，誤差が約 5%程度であり，高い予測精度であることが分かる。また，決定係数も高い。

4.2 重要な説明変数・特徴量

上記 3 つのモデルにおいて，機械学習ライブラリ scikit-learn を用いて，全 137 個の説明変数・特徴量の重要な上位 10 個を求め，その中で共通する説明変数・特徴量を表 1 にまとめた。相関分析と同様，「1 株益」や「1 株当たり純資産」が重要な説明変数・特徴量であることが分かった。これらは，代表的な理論株価でも利用されており，株価を説明する際，重要な特徴量であることが分かった。

表 1　各モデルに共通する重要な説明変数・特徴量

3 つ	連結・1 株益(修正，最新) 連結・1 株当たり純資産 (修正，最新)
2 つ	上場区分 業種 発行済株式数(対数) 連結・1 株配当(修正，最新) 連結・1 株配当(修正，1 期前) 連結・1 株配当(修正，4 期前) 連結・剰余金の配当 連結・1 株当たり純資産 (修正，5 期前) 連結・売上高原価率(%)

5. まとめと今後の課題

以上のように，会社四季報のデータを，Kaggle で利用されている手法を活用し，理論株価を導出した。指摘されているように，本稿においても分析アルゴリズムよりも説明変数・特徴量を工夫することにより，大きく精度が改善することが確かめられた。

本分析は，あるモデルの結果を示したものであり，より精度の高いモデルもあると考えられ，試行錯誤，様々なデータ分析を行うことが重要である。

実際，Kaggle は，データ分析を行うものにとって，失敗も含め，様々な情報が蓄積されてきており，データ分析を行う上で，大変有益な情報が多い。

本稿での分析における欠点として，例えばデータに関する問題がある。

(1) 第 2 集(春号)から第 3 集(夏号)へのデータの更新・変更があまりない。そのため，数字の修正があったとしても，カテゴリーが変わるほど影響がない。

(2) 例えば第 2 集(春号)では欠損値であったものが，第 3 集(夏号)となり，データが入力されることがある。この変化はファンダメンタルの変化ではなく，時間が経過し，データが入力されたに過ぎない。

また，残された課題として，

(1) 会社四季報編集部による業績予想数字と企業評価の手がかりとなるような材料についてのコメントには変更があるため，これに関するテキストマイニングを行うとより，今後の予測精度向上となるかもしれない[8]。

(2) 実際にこのような分析結果を利用する際には，本稿では，2019 年第 2 集(春号)から第 3 集(夏号)のみの予想を行なったが，データの期間を遡り検証を重ねていくことが必要である。

注意事項

本稿における理論株価は，あくまで「会社四季報」から得られる財務・業績数値等から独自に算出した「理論的な株価水準」を計算したものであり，その正確性，実現性を保証するものではない。

[8] 会社四季報のコメント欄にある言葉がポジティブかネガティブかを判断し，銘柄評価に活用しているファンドもある[10]。

参考文献

[1] https://shikiho.jp/
[2] https://www.kaggle.com/c/mercari-price-suggestion-challenge/
[3] https://www.kaggle.com/c/recruit-restaurant-visitor-forecasting
[4] https://www.kaggle.com/c/two-sigma-financial-news
[5] 片山晃，小松原周「勝つ投資 負けない投資」クロスメディア・パブリッシング，2015 年
[6] https://drive.google.com/file/d/1G1uJoC2umBh3zsg0oGH2dKP4nfpmbe3L/view
[7] 門脇大輔，阪田隆司，保坂桂佑，平松雄司「Kaggle で勝つデータ分析の技術」技術評論社，2019 年
[8] Alice Zheng,Amanda Casari「機械学習のための特徴量エンジニアリング―その原理と Python による実践」オライリージャパン，2019 年
[9] https://www.kdnuggets.com/2016/03/xgboost-implementing-winningest-kaggle-algorithm-spark-flink.html
[10] https://www.gsam.com/content/gsam/jpn/ja/gsitm/resources/investment-ideas/gs-bdsj.html#section-background_aa2c_background_sectionheaderwithtab
[11] https://str.toyokeizai.net/files/user/dl/magazine/risc_riron.pdf

付録　広く知られている理論株価

	東洋経済新報社「会社四季報プロ 500」[9]	ダイヤモンド社「ダイヤモンド ZAi(ザイ)」[10]
理論株価	理論株価＝直近期 BPS ＋将来の 1 株当たりの超過利益の現在価値	理論株価＝資産価値＋利益価値＋成長価値
詳細	将来の 1 株当たりの超過利益の現在価値＝（今期予想 EPS－最近 BPS×株主資本コスト）/（1＋株主資本コスト)＋（来期予想 EPS－今期予想 BPS×株主資本コスト）/ (1+株主資本コスト)^2 ＋　以下 n 期　まで 株主資本コスト＝長期国債（10 年債）利回り ＋β 値×株式リスクプレミアム EPS = 1 株当たり利益 BPS = 1 株当たり純資産 注　特殊要因による過大な理論株価を避けるため，時価に対する倍率を最大 3 倍に制限している。	資産価値：1 株あたりの株主純資産(BPS) 利益価値：今期の予想 1 株益(EPS)×年数 成長価値：売上高成長率から算出した，各年の 1 株益成長分の合計
利用データ	純資産（前期末，最新） 発行済株式数（最新） 予想利益(今期，来期） 長期国債（10 年債）利回り 業種	純資産（前期末） 発行済株式数（最新） 前期売上高（前期末） 今期予想売上高(最新） 予想利益(最新） 業種
MAPE	75.07	83.79

[9] 「会社四季報プロ 500」は編集部が選んだ「会社四季報」にある 500 銘柄について，図を多く入れることで，よりビジュアル的に分かりやすく解説した投資情報誌である。この理論株価に関しては，[11]で説明されている。

[10] 四半期毎に，別冊付録として理論株価を掲載している。

機械学習によるテクニカル分析の影響の調査

Analysis of effect by technical analysis using machine learning

片寄　諒亮 [1*]　吉岡　真治 [1,2]
Katayose Ryosuke[1]　Masaharu Yoshioka[1,2]

[1] 北海道大学大学院情報科学院
[1] Graduate School of Information Science and Tehnology, Hokkaido University
[2] 北海道大学大学院情報科学研究院
[2] Faculty of Information Science and Tehnology, Hokkaido University

Abstract: In recent years, stock prices have been predicted in various forms such as technical analysis and fundamental analysis using time series data and financial indicators, and text analysis using news information. In particular, some use text mining to predict whether stock prices will rise or fall from text information, but useful text information does not appear frequently on all stocks. Considering the increase in algorithmic trading using technical analysis, analysis that relies solely on textual information is not appropriate because it does not take into account the impact of such trading. Therefore, in this study, first, stocks that are easily affected by technical analysis were ranked by using machine learning to raise and lower stock prices using indicators that are often used in technical analysis techniques. Moreover, when the analysis did not go well, we analyzed what kind of events occurred and investigated how technical analysis affects the stock price.

1　はじめに

1.1　背景

近年、時系列データや金融指標を用いたテクニカル分析やファンダメンタル分析、さらにニュース情報を用いたテキスト分析など様々な形で株価の予測が行われている。特にテキストマイニングを使ってテキスト情報から株価の騰落を予測するものもあるが、全ての銘柄に対して頻繁に有用なテキスト情報が出るわけではない。また、テクニカル分析などを用いたアルゴリズム取引が増えていることを考慮すると、テキスト情報だけに頼った分析では、このような取引の影響が考慮されないため、適切な分析にならないことが想定される。そこで本研究では、まずテクニカル分析の手法で良く用いる指標を複数用いて、株価の騰落を機械学習によって行うことでテクニカル分析の影響を受けやすい (予想があたりやすい) 株式のランク付けや、その分析がうまく行かないときに、どのような事象が発生しているのかを分析することで、テクニカル分析が株価にどのような影響を与えるかについての調査を行なった。

*連絡先：北海道大学大学院情報科学院
〒 060-0814　北海道札幌市北区北 14 条西 9 丁目
lk@eis.hokudai.ac.jp

1.2　株価の変動要因

現在株価の変動要因と考えられているものは大きく４つある。

- ファンダメンタルズ (国や企業などの経済活動の状況を表すもの)
- テクニカル (株価の変動パターンのような過去の時系列的動きから作成されるもの)
- センチメント (相場状況や様々な情報に対する投資家心理)
- アノマリー (理論で説明できない規則性)

主にこれらの４つの変動要因により株価は動いていると考えられる。この変動要因を元に株価予測を行なう研究が多く行なわれているが、背景でも述べたように現状ではニュースの分析からの株価予測 [1] や、センチメントに対する分析 [2] など、その１つの要因からの分析によって株価を予測するものが多い。しかし背景でも述べたようにテキストマイニングを元にした予想を例にすると実際にはある銘柄に対して、毎日投資家に影響を及ぼすようなニュースが出ることは少ない。

実際の株価の変動要因は上で挙げたが、参照とするデータが数値データで表せるのかどうかという観点か

ら変動要因を大きく 2 つに分けた [3]。数値データ以外のデータというのは全ての銘柄に対して、毎日有用であるようなものではないと考えられることから、数値データを主要因、それ以外を付属要因とした。イメージとしては上の変動要因としたファンダメンタルズが、全体の景気動向を元にした変動と、個別銘柄に関する基本的情報の 2 種類に分かれているようなものになる。

- 主要因 (数値データを元としたもの)
    1. 企業や国全体の景気動向
    2. テクニカル分析

- 付属要因 (数値データ以外を元としたもの)
    1. 個別銘柄に関する基本的情報 (業績の情報や不祥事などのニュース)
    2. センチメントに関するもの (情報に対しての投資家の反応)
    3. アノマリー (理論で説明できない規則性)

実際にこのように数値データとそれ以外に分けた研究 [4] は過去にされており、この研究では主に数値データの方の要因に加えてニュースで報道される社会的イベントを考慮した株価予測を行なっている。

本研究では、この過去の研究の内容から株価予測について、次のような新たな仮説を考えた。

仮説 株価変動の主要因としたものの影響を調べることで、今回付属要因として位置付けたものの要因が実際の株価変動要因になっている部分を調べることができるのではないか。

詳しく説明すると、企業や国全体の景気動向、そしてテクニカル分析から株価を予想した際にある程度の予想ができるモデルができたとすると、それが適用できない範囲、すなわち予測が外れる時期であったり、条件などを洗い出すことができれば、データ量が少ない業績に関するニュースも、研究が多く行われているテキストマイニングによる株価予測もこれに適応しやすくなるのではないかと考えた。

その上で、以下のような検証を行うことにより、本仮説の有用性が検討できると考えた。

1. 企業や国全体の景気動向は個別銘柄に対して影響を与えているのかを、日経平均との相関性を調べることにより検証する。
2. テクニカル分析においても有名な手法をモデル化することによって、それがうまく当てはまるのかを検証する。

本論文では、1 と 2 についての検証方法とそれを行なった結果と考察を述べる。

2 テクニカル分析モデル

2.1 作成したモデルについて

今回提案する手法はシステムトレードにおいてよく使われているテクニカル分析の特徴量を用いて株価の予測を行うモデル作成した。今回使用したデータは 2018 年時点で日経 225 に入っている 225 銘柄と日経 225 の株価と連動して動く日経 225 連動型上場投資信託の合わせて 226 銘柄についての 2009 年から 2018 年の 10 年間の株価データである。2009 年以降に上場した企業についてはこれよりもデータが少ないこともある。これらの数値データをもとに様々なテクニカル指標を算出した。また作成したモデルは 226 銘柄すべての 2009 年以降のデータから 2017 年までのデータをトレーニングデータとして使用して作成し、これを 2018 年のデータをテストデータとして株価の予測を行なった。今回、用いたテクニカル分析の特徴量として、変動率に加えて、システムトレードにおいてよく使われているテクニカル指標の中から、移動平均線乖離率、RSI、MACD、ATR、モメンタム、パラボリック、DMI、ストキャスティクスを使用した。またこれらの指標から 15 個の特徴量を作成し、翌日の変動率が上がるか下がるかを予測した。この特徴量には次に説明するテクニカル指標によって判断したシグナルの有無を入れている。また、Python の機械学習ライブラリである scikit-learn を使用し、機械学習の一種である randomforest を使用してモデルの作成を行い予測を行なっている。

2.2 テクニカル指標

今回有名なテクニカル分析をモデル化するにあたって、機械学習に説明変数として用いるテクニカル分析の手法を説明する。移動平均線乖離率、RSI、MACD、パラボリック、DMI、ストキャスティクスについてはマネックス証券のテクニカル指標一覧 [5] にもあるがいくつかを説明する。また [5] では紹介されていない ATR、モメンタムについての説明も行う。

ATR(Average True Range) ATR とは TR の移動平均をとったもので、TR は当日の高値から当日の安値を引いたもの、当日高値から当日安値を引いたもの、当日高値から前日終値を引いたものそれぞれ以下の式で表すことができる。

$$ATR = Avg(TR)$$

$$TR = Max \begin{pmatrix} T_{high} - T_{low} \\ T_{high} - Y_{close} \\ Y_{close} - T_{low} \end{pmatrix}$$

ただし、T_{high}：当日高値、T_{low}：当日安値、Y_{close}：前日終値これによって株価の動く大きさ、すなわちボラティリティが高いか低いかを表すことができる。つまり ATR は株価の上昇、下降それぞれの局面における勢いの強さを判断することができる指標であり、ATR が上昇している時には株価のトレンドが持続する可能性が高く、ATR が上昇から下降に転ずる時には株価のトレンドが弱まり、反転する可能性があることを示している。

DMI(Directional Movement Index) この指標の特徴は相場の強弱を可視化し、価格の変動幅（ボラティリティ）からトレンド分析を行える部分にあります。DMI は通常、3 つの指標 (DI^+, DI^-, ADX) から総合的に売買の判断をします。

$$DM^+ = T_{high} - Y_{high},\ DM^- = Y_{low} - T_{low}$$

ただしそれぞれ負の値になる場合や DM^+ であれば DM^- を下回った場合、DM^+ は 0 となる。逆も同様である。この値により、3 つの指標が以下のように導かれる。また ADX が上昇している際に DM^+ が DM^- を上に突き抜けた場合に買いシグナル、逆の場合に売りシグナルとなる。

$$DI^+ = \frac{\sum^n DM^+}{\sum^n TR} \times 100,\ DI^- = \frac{\sum^n DM^-}{\sum^n TR} \times 100$$

$$ADX = \frac{1}{n} \sum^n \frac{|DI^+ - DI^-|}{DI^+ + DI^-}$$

モメンタム (Momentum) 10 日間のモメンタムは、当日の終値から 10 日前の終値を引くだけで求めることができ、これによって相場の短期的な動きを判断する指標である。すなわち、相場が上昇している時の勢いが弱くなってきているのか、また相場が下降している時の勢いが強くなってきているのかを捉えることができる。

移動平均線 (Moving Average) 移動平均線とは一定期間の平均終値を計算し、それぞれの結果を線でつないだものである。例えば、5 日移動平均値は 5 日間の平均価格であり日々の移動平均値を線でつないだものが移動平均線となる。
これを用いて、テクニカル分析でよく行なわれているのは、ゴールデンクロスとデッドクロスを利用したものである。ゴールデンクロスとは短期の移動平均線が長期の移動平均線を下から上へ追い越した場面のことを言う。デッドクロスはその逆である。一般的にはゴールデンクロスが起こった際には今後の株価は上がっていくことが予想される。デッドクロスはその逆で起これば、それ以降の株価は下がっていくことが予想される。

RSI(Relative Strength Index) 相場において買われすぎか売られすぎかを値上がり率の比率から求める有名な指標である。一般的には 30%を下回ると売られ過ぎていると言うことで株価は好転することが多く、逆に 70%を越えると買われすぎだと判断する。

$$RSI = \frac{Avg_{sum}}{Avg_{gain} + Avg_{loss}} \times 100$$

ただし、
Avg_{sum}：一定期間の値上がり幅の平均、
Avg_{gain}：一定期間の値上がり率の平均、
Avg_{loss}：一定期間の値下がり率の平均

MACD(Moving Average Convergence Divergence) 短期の移動平均線と長期の移動平均線の差を MACD と言う。この MACD の移動平均線をとったものがシグナルと呼ばれ、MACD がシグナルを上に突き抜けたら「買い」と判断する。すなわち今後の株価は上がっていくと予想できる。一方、MACD がシグナルを下に突き抜けたら「売り」と判断し、今後の株価は下がっていくと予想する。移動平均線のゴールデンクロスやデッドクロスによるシグナルに先行する形で、MACD の売買シグナルが出現する特徴がある。
特にここででた短期の移動平均線と長期の移動平均線の差を乖離といい、その割合をとったものを乖離率という。

3 調査結果と考察

3.1 主要指標との連動性

日経 225 の変動率それぞれの個別銘柄の変動率データを使用し、相関係数を計算した。ここでの変動率は前日の終値と当日の終値の差の割合を表している。実際には株式の分割や併合を行なった銘柄については相関係数が極端に低くなってしまったため、それに該当する 51 銘柄を除いた 174 銘柄についての検証を行なった結果を表 1,2 に示す。

相関関係の高い銘柄に関してはあまりニュースにもならず手堅い企業のイメージがある企業が名を連ねた。一方で相関関係が低い銘柄に関してはデータ改ざんにより株価を大きく下げた時期があった神戸製鉄所や比較的新しい事業のニュースが多く出る IT 企業が入った。

この 2 つの対象の銘柄の中でも最も日経 225 との変動率に差があったのは 1333(マルハニチロ) の 2 月 6 日

表 1: 日経 225 との相関関係が高い 5 銘柄

銘柄名	相関係数
8015(豊田通商)	0.75100
6367(ダイキン工業)	0.74736
7267(ホンダ)	0.71834
5333(日本ガイシ)	0.71530
6503(三菱電機)	0.71146

表 2: 日経 225 との相関関係が低い 5 銘柄

銘柄名	相関係数
1333(マルハニチロ)	0.30097
9503(関西電力)	0.34878
5406(神戸製鋼所)	0.37308
4689(Z ホールディングス[1])	0.37921
2432(ディーエヌエー)	0.38702

であった。図 1 についてはもっとも変動率の差が大きかった 2 月 6 日を含んだ 2 月における日経 225、豊田通商、マルハニチロの 3 銘柄の株価変動率を表したものである。この日付について調査したところ、前日の 2 月 5 日に決算予想の発表があり、予想業績を上方修正したことにより個別銘柄の株価は上昇したことが考えられる。一方で、日経平均は同じく前日にアメリカの株価が過去最大の下げ幅を記録したことから、その影響を受け、大幅に値を下げたことが考えられる。

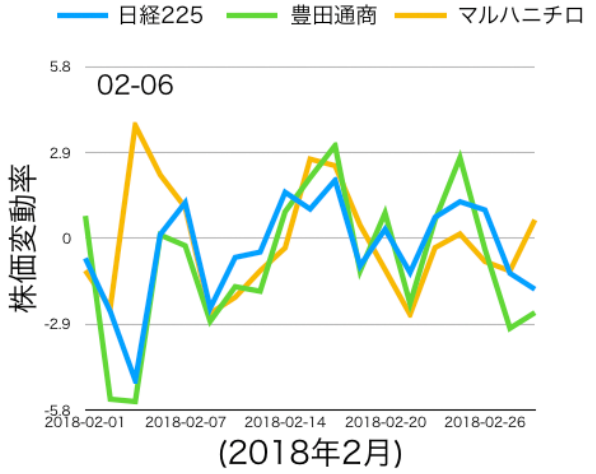


図 1: 相関係数を元にした 2018 年の株価変動率の推移 (2 月)

この結果からもわかるように独自の情報、すなわち決算であったり不祥事のニュースがある場合では連動性が下がる可能性があることが確認できた。このように独自情報がある日付については、センチメント解析や個別銘柄に関するニュースについてのテキストマイニングなど、他の情報から予測を行う必要がある。

しかし、このような独自の情報は決して多くはないため、これらによる変動が起きた時期を見つけることは、個別要因の影響を分析する上で重要な事例となる。

3.2 テクニカル分析

今回の調査では日経 225 において使われている銘柄についてすべての銘柄について、2 日前と前日の離散化した株価変動（上昇、下降）と 2.2 節で述べた指標に基づく売り買いのシグナルが発生したかどうか（買いシグナル、シグナルなし、売りシグナル）をフィーチャーとして、当日の株価の変動（上昇、下降）を当てるという二値分類の問題としてモデル化し、機械学習による予測を行なった。具体的には、scikit-learn[6] のランダムフォレストを用い、訓練用のデータとしては、2009 年から 2017 年の株価に基づいたデータを利用し、2018 年の株価の変動を当てるというタスクを実施した。ただし、日経 225 の銘柄にも変動があるため、トレーニングデータについては、日経 225 に登録された期間のものに限定することとした。この結果、精度の良かった 5 銘柄と悪かった 5 銘柄を表 3,4 に示す。た

表 3: 機械学習に基づく予測上位 5 銘柄

銘柄名	予測精度 (%)
6773(パイオニア)	62.832
6752(パナソニック)	61.504
7004(日立造船)	61.062
4502(武田薬品工業)	59.735
4689(Z ホールディングス)	59.292

表 4: 機械学習に基づく予測下位 5 銘柄

銘柄名	予測精度 (%)
8028(ファミリーマート)	37.168
9735(セコム)	42.035
4506(大日本住友製薬)	43.363
4902(コニカミノルタ)	43.805
9613(NTT データ)	44.248

だし、全体の結果としては 51.750%となり、使用したテクニカル指標の良し悪しや予想する時期を長期でみるのか短期でみるのかなどについて、再検討が必要であると考えている。また予測精度が低かった、8028(ファミリーマート) と 9735(セコム)、4506(大日本住友製薬)

[1] 旧ヤフー

については 2018 年 1 年間での株価の上昇率が高く、急激な上昇もいくつかみられる。このような銘柄は、業績その他の材料が多い銘柄だと考えられる。これに対し、7004(日立造船)、6752(パナソニック)、4502(武田薬品工業) は、全体的に株価の下落率が高い銘柄になっている。

4 まとめと今後の課題

実験を行う前の予想としてはシステムトレードを行う人が少なからずいることからもテクニカル分析によるシグナルが出た後には、ある程度の割合でシグナルなどに沿って株価が変動するものだと考えていた。現段階ではそれをうまく示すことができる有用なモデルはできなかったものの、各銘柄においてそれぞれの特徴があることや一概にテクニカルな分析が全てにおいては有用ではないということが議論できた。

今回のモデルの予測があまり良くなかった要因として考えられるものは以下に示されるように 2018 年以外の株価変動と 2018 年の株価の変動率が違うということが考えられる。以下の表 5 は今回用いた全銘柄の変動回数を表している。あるいは急激な値下がりについてはリーマンショックなどでデータがそれなりに多く存在するのに対し、急激な値上がりというのは、あまりデータがないことも一因だと考えられる。

表 5: 全 226 銘柄の株価変動回数

	上昇	下落
2018 年以外の株価変動	253681 回 (52.03%)	233932 回 (47.97%)
2018 年の株価変動	24356 回 (47.69%)	26720 回 (52.31%)

今後の課題としては予測に使う手法や特徴量の入れ方を見直すことや使うテクニカル指標の種類を変えることで、今回株価変動の主要因として考えてきた数値データによる予測の精度を上げるということが 1 番に考えられる。特に近年 LSTM[7] や CNN[8] を予測に用いる手法が成果を出していることから今回作成したテクニカル指標のデータを用いることで予測精度の改善に繋げていきたい。これができた上で、これらの予測をさらに補正していけるように、今回は付属要因として考えた、数多くは出ない企業の決算情報であったり、ニュースのテキスト記事などを加えて行くことにつなげていくことが目標となる。

5 参考文献

[1] 敷地琢也, 和泉潔. 複数のテキスト情報を用いた株式市場動向の分析. 人工知能学会研究会資料 (第 13 回金融情報学研究会).SIG-FIN-013-09.（2014）.

[2] 河合継, 新田翔, 大川尭郁, 西山昇. 感情によるマルチモーダル AI を利用した IPO 株価推定. 人工知能学会研究会資料 (第 20 回金融情報学研究会).SIG-FIN-020-03.（2018）.

[3] 三菱 UFJ トラスト投資工学研究所 [編]. 実践金融データサイエンス. 日本経済新聞出版社. (2018) .

[4] 中菅章浩, 関和広, 上原邦昭. ニュース記事で報道される社会的イベントを考慮した株価動向予測の補正. 人工知能学会研究会資料 (第 10 回金融情報学研究会).SIG-FIN-010-03.（2013）.

[5] マネックス証券 テクニカル指標一覧
https://info.monex.co.jp/technical-analysis/indicators/

[6]scikit-learn
https://scikit-learn.org/

[7]David M.Q. Nelson,Adriano C.M.Pereira,Renato A. de Oliveira.Stock market's price movement prediction with LSTM neural networks.International Joint Conference on Neural Networks (IJCNN). (2017)

[8] 宮崎邦洋, 松尾豊. 深層学習を用いた株価予測の分析. 人工知能学会全国大会論文集 (第 31 回人工知能学会全国大会).2D3-OS-19a-3. (2017)

多相アテンションLSTMを用いた財務時系列データの予測

Forecasting Corporate Financial Time Series using Multi-phase Attention Recurrent Neural Networks

吉見 脩平[1*]　江口 浩二[2†]　金京 拓司[3]　羽森 茂之[3]
Shuhei Yoshimi[1]　Koji Eguchi[2]　Takuji Kinkyo[3]　Shigeyuki Hamori[3]

[1] 神戸大学大学院システム情報学研究科
[1] Graduated School of System Informatics, Kobe University
[2] 広島大学情報科学部
[2] School of Informatics and Data Science, Hiroshima University
[3] 神戸大学大学院経済学研究科
[3] Graduated School of Economics, Kobe University

Abstract: Recently, attention-based RNNs have been studied to represent multivariate temporal or *spatio-temporal* structure underlying multivariate time series. One recent study has achieved improved performance by employing attention structure that simultaneously capture the *spatial* relationships among multivariate time series and the temporal structure of those time series. That method assumes a single time-series sample of multivariate explanatory variables, and thus, no prediction method was designed for multiple time-series samples of multivariate explanatory variables. Moreover, such previous studies have not explored on financial time series incorporating macroeconomic time series, such as Gross Domestic Product (GDP) and stock market indexes, to our knowledge. Also, no neural network structure has been designed for focusing a specific industry. We aim in this paper to achieve effective forecasting of corporate financial time series from multiple time-series samples of multivariate explanatory variables. We propose a new industry specific model that appropriately captures corporate financial time series, incorporating the industry trends and macroeconomic time series as side information. We demonstrate the performance of our model through experiments with Japanese corporate financial time series in the task of predicting the return on assets (ROA) for each company.

1　はじめに

近年，情報技術の発達に伴い，日々多くの情報が収集，蓄積されている．その分野は多岐にわたり，枚挙に暇がないが，その一つに，経済・金融分野における企業財務時系列データが挙げられる．これらを活かして新たな知見を得ることについて当該分野において関心が高まりつつある．企業財務時系列データは，企業の経営状況や業界の動向，さらには社会の景況感などの影響を受けた複雑な時系列データとなっており，その予測のためには複数の時系列の分析を必要とする．従来の時系列分析や最近の深層学習の研究において，時系列予測の問題に広く取り組まれてきたものの，企業の財務時系列などの複雑な多変量時系列に関しては新たな研究の余地が十分にある．広く認知された時系列予測手法の一つに自己回帰統合移動平均モデル（Autoregressive integrated moving average: ARIMA）があり，時系列における時間的パターンの 1 つの側面の分析に焦点が当てられているものの，多変量時系列の的確な長期予測を達成することは容易ではないとされる [1]．再帰型ニューラルネットワーク（Recurrent neural network: RNN），および，RNN の欠点を改善して長期にわたった時間依存性を考慮できる長・短期記憶モデル（Long short-term memory: LSTM）は，株価予測などの金融時系列分野における有用な手法の一つとして注目されつつある [8]．しかしながら，多変量時系列の適格な長期予測を達成するには解決すべき課題が依然として少なくない．具体的には，多変量の時系列を考慮するこ

*連絡先：神戸大学大学院システム情報学研究科
〒 657-8501 兵庫県神戸市灘区六甲台町 1-1
E-mail:shuhei@cs25.scitec.kobe-u.ac.jp
†連絡先：広島大学情報科学部
〒 739-8527 広島県東広島市鏡山 1-4-1
E-mail:eguchi@acm.org

とにより，一部の説明変数時系列が目的変数時系列の予測への寄与が小さいときに，その予測性能に悪影響を与え得る．このような問題意識のもと，複数時系列に対して予測に関する重要度の分布を学習して複数時系列の特徴表現能力をさらに強化する目的で，アテンションメカニズムに基づく RNN を用いて，時系列間の相関（以下，空間的相関）を捉えつつ時間的関連性を抽出する予測モデルが提案された [2]．また，Qin ら [11] と Yuxuan ら [9] は，アテンションメカニズムとエンコーダ・デコーダモデルを組み合わせて，単一ステップおよび短期時系列予測での予測性能を向上させた．さらに，Liu ら [10] は，多段階のアテンションメカニズムを介して空間的相関へのアテンションを強化しつつ，目的変数時系列の埋め込みを介して時間的関連性へのアテンションを強化するような，2 段 2 相（dual-stage two-phase: DSTP）のアテンションメカニズムに基づく RNN モデルを開発した．しかしながら，それらの研究では多変量説明変数を伴う複数の時系列サンプルが与えられた状況での分析と予測に関して報告されていない．また，国内総生産（GDP）や日経平均株価などのマクロ経済時系列を組み込んだり，業種ごとの傾向に着目するような RNN モデルによる財務時系列分析の研究は，著者の知る限りない．本論文では，企業財務時系列の予測に対する有用な手法を確立するため，アテンションメカニズムに基づく RNN モデルを拡張し，多変量説明変数を伴う複数時系列サンプルに対する的確な予測を目指す．また，マクロ経済時系列を適切に捉えたモデル化を行う．さらに，特定の業種の動向を適切に反映したモデルを提案する．日本の企業財務時系列を用いて各企業の総資本経常利益率（ROA）を予測する実験を行い，提案モデルが有効に機能することを示す．

2　関連研究

本章は次の 3 つの話題からなる．最初の話題としては，時系列データを予測するための基本的な深層学習手法の 1 つとして RNN を概説する．次に 2 番目の話題として，RNN における長期依存性を考慮できない問題に対処すべく拡張された LSTM を紹介する．さらに 3 番目の話題として，近年注目を集めつつあるアテンションメカニズムを取りあげる．

2.1　再帰型ニューラルネットワーク（RNN）

系列データは，サンプルの順序が重要であるようなデータの総称である．特に，その順序が時間に基づいている場合は時系列データと呼ばれる．過去のサンプルと現在のサンプルとの間の依存関係をモデル化する研究が，時系列予測の問題に対して取り組まれてきた．なかでも広く受け入れられた手法の 1 つに再帰型ニューラルネットワーク（Recurrent Nural Network: RNN）があり，順伝播型ニューラルネットワークを系列データの処理のために拡張したモデルとして位置づけられる [3]．ここで，RNN は各時間 t の入力系列 $\mathbf{x}_t$ を受け取り，時間 t で 1 つの出力系列 $\mathbf{y}_t$ を返すように定式化する[1]．時間 $t=n$ において出力系列 $\mathbf{y}_n$ が入力系列 $(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n)$ から正常に生成されたと仮定できる．これは RNN が「リターンパス」と呼ばれる有向閉ループを使用しているためである．この構造により，時間情報の保存と更新が可能になる．図 1 に，RNN の構造と，時間次元で展開された構造を示す．図中で，$\mathbf{x}_t$ は入力，$\mathbf{h}_t$ は隠れ状態，$\mathbf{y}_t$ は出力を示し，それぞれの層を入力層，中間層，出力層と呼ぶ．$W^{(in)}$ は入力層から中間層への重み，$W^{(out)}$ は中間層から出力層への重みを示している．RNN では時間 $t-1$ の隠れ状態から時間 t の隠れ状態への接続が想定され，時間 $t-1$ における任意の隠れ状態と時間 t の任意の隠れ状態の間に重み $w \in W$ が割り当てられる．図 2 に RNN の中間層の 1 つのユニットを示す．

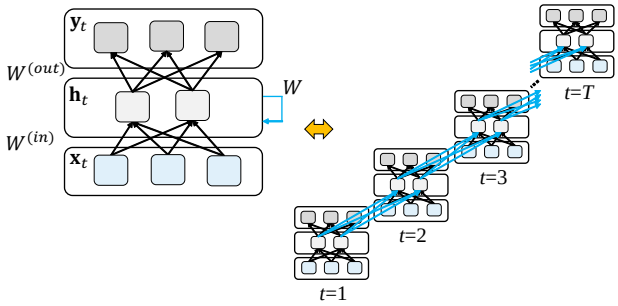


図 1: RNN の構造.

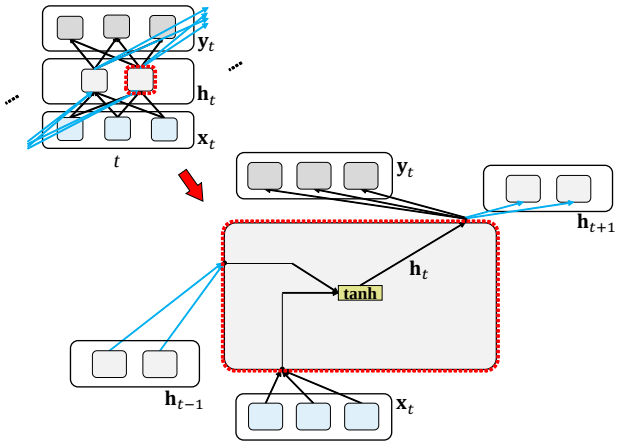


図 2: RNN の中間層のユニット.

[1]後述する提案モデルでは，各時間 t の多変量説明変数系列 $\mathbf{X}_t$ を受け取り，時間 $t+1$ で 1 つの目的変数系列 $\mathbf{y}_{t+1}$ を返すように定式化する．

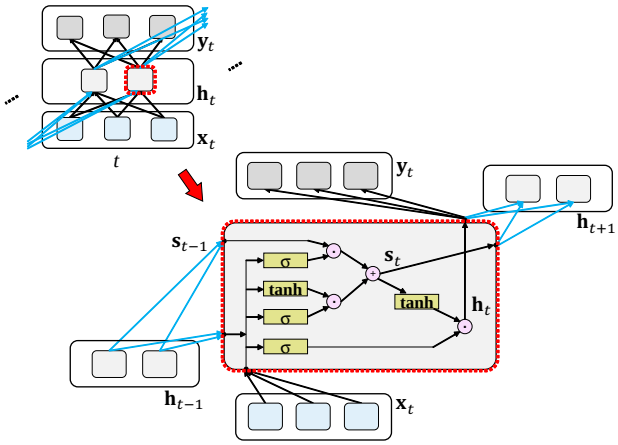


図 3: LSTM の中間層のユニット．

2.2 長・短期記憶モデル（LSTM）

RNN により系列データのコンテキストを捉えることができるが，そのとき，現在の時刻からどれだけ過去の入力を反映して出力を予測するかを考えることが重要である．RNN の未知パラメータを学習する際に目的関数の勾配を用いて更新を繰り返すのが基本となるが，一定数以上の繰り返しで勾配が消失（または発散）することが知られており，勾配消失問題と呼ばれる [4, 5]．この問題に対処するために，長期記憶を実現する長・短期記憶モデル（Long Short Term Memory: LSTM）が提案されている [6]．図 3 に LSTM の中間層の 1 つのユニットを示す．LSTM では中間層の各ユニットにおいて情報を記憶するためのセル（$\mathbf{s}_t$）と情報の流れを制御するためのゲートが配置され，その他の構造は基本的に RNN と同じである．一般に LSTM は RNN の一種と位置づけられる．

2.3 アテンションメカニズム

アテンションメカニズムは LSTM に使用され，その有効性が示されている [11, 10]．LSTM に基づく時系列予測において，アテンションメカニズムにより多変量時系列の間の相関とそれらの時間的構造を同時に捉えることができる．本節の残りでは，文献 [11, 10] において LSTM のために開発されたアテンションメカニズムを概説する．

空間的アテンション LSTM　空間的アテンションメカニズム（Spatial attention mechanism）は，多変量時系列の間の相関（以下，空間的相関）を抽出することを目的としている [11, 10]．時刻 t までに観測された多変量時系列（長さ T，変量数 n）における k 番目の変量の系列を $\mathbf{x}_t^k = (x_{t-T+1}^k, \cdots, x_t^k) \in \mathbb{R}^T$ とするとき，次のようなアテンションメカニズムを考える．

$$a_t^k = \mathbf{v}_a^{\mathrm{T}} \tanh(W_a[\mathbf{h}_{t-1}; \mathbf{s}_{t-1}] + U_s \mathbf{x}_t^k + \mathbf{b}_a) \quad (1)$$

$$\alpha_t^k = \frac{\exp(a_t^k)}{\sum_{j=1}^n \exp(a_t^j)} \quad (2)$$

ここで，$[\cdot;\cdot]$ は連接演算を示し，$\mathbf{h}_{t-1} \in \mathbb{R}^m$ および $\mathbf{s}_{t-1} \in \mathbb{R}^m$ はそれぞれ時間 $t-1$ に対応する LSTM ユニットの隠れ状態およびセル状態を示し，m は隠れ状態数を示す．また，$\mathbf{v}_a, \mathbf{b}_a \in \mathbb{R}^T, W_a \in \mathbb{R}^{T \times 2m}, U_a \in \mathbb{R}^{T \times T}$ は未知パラメーターを示す．このように時間 $t-1$ の隠れ状態と時間 t までの多変量入力時系列によって，時間 t における各変量の重要性の度合い（以下，空間的アテンション重み）$\boldsymbol{\alpha}_t = (\alpha_t^1, \cdots, \alpha_t^n)$ を求める．この重みを用いて，時間 t の時点での多変量入力 $\mathbf{x}_t = (x_t^1, \cdots, x_t^n)$ を次式のように重みづける．

$$\tilde{\mathbf{x}}_t = (\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \cdots, \alpha_t^n x_t^n)^{\mathrm{T}} \quad (3)$$

以上に述べたような空間的アテンション LSTM ユニットを $f_{spatial}$ で表すと次式を得る．

$$(\mathbf{h}_t, \mathbf{s}_t) = f_{spatial}(\mathbf{h}_{t-1}, \mathbf{s}_{t-1}, \tilde{\mathbf{x}}_t) \quad (4)$$

時間的アテンション LSTM　時間的アテンションメカニズム（Temporal attention mechanism）は，前項の空間的アテンションの時間依存性を捉えることを目的としている [11, 10]．長さ T の時間ウィンドウにおける時空間構造を抽出するために，前項で述べた多変量時系列の空間的相関を反映させた隠れ状態ベクトルを用いる．固定幅ウィンドウで時間的関連性を抽出するだけでは十分ではないため，隠れ状態を選択するためのアテンションメカニズムを適用する．具体的には，予測対象の目的変数に関連性が高い隠れ状態を選択することにより長期依存性を反映させる．アテンションメカニズムにより，i 番目の隠れ状態に関して各時間の重要性の度合い（以下，時間的アテンション重み）$\boldsymbol{\beta}_i = (\beta_i^1, \cdots, \beta_i^T)$ を次のように求める．

$$b_t^i = \mathbf{v}_b^{\mathrm{T}} \tanh(W_b[\mathbf{d}_{t-1}; \mathbf{s}'_{t-1}] + U_b \mathbf{h}_i + \mathbf{b}_b) \quad (5)$$

$$\beta_t^i = \frac{\exp(b_t^i)}{\sum_{j=1}^T \exp(b_t^j)} \quad (6)$$

ここで，$\mathbf{h}_i$ は，前項で述べた空間的相関を表す隠れ状態の i 番目のベクトルを示す．$\mathbf{d}_{t-1} \in \mathbb{R}^p$ および $\mathbf{s}'_{t-1} \in \mathbb{R}^p$ はそれぞれ時間 $t-1$ に対応する LSTM ユニットの隠れ状態およびセル状態を示し，p は隠れ状態数を示す．$\mathbf{v}_b, \mathbf{b}_b \in \mathbb{R}^p, W_b \in \mathbb{R}^{p \times 2p}, U_b \in \mathbb{R}^{p \times m}$ は未知パラメータを示す．次に，コンテキストベクトル $\mathbf{c}_t$ を次のように定義する．

$$\mathbf{c}_t = \sum_{j=1}^T \beta_t^j \mathbf{h}_j \quad (7)$$

コンテキストベクトル $\mathbf{c}_t$ はすべての隠れ状態の重み付き和を表し，時間ウィンドウ内における時間的関連性を表す．

このようにして求めたコンテキストベクトル $\mathbf{c}_t$ と目的変数 y_t を次式のようにして組み合わせる．

$$\tilde{y}_t = \tilde{\mathbf{w}}^{\mathrm{T}}[y_t; \mathbf{c}_t] + \tilde{b} \tag{8}$$

ここで，$\tilde{\mathbf{w}} \in \mathbb{R}^{m+1}$ および $\tilde{b} \in \mathbb{R}$ は連接ベクトルを目的変数に写像するパラメータである．目的変数時系列をコンテキストベクトルに対応付けることにより時間的関連性を管理しやすくなり，その結果を使用して隠れ状態を更新することができる．以上に述べたような LSTM ユニットを $f_{temporal}$ で表すと次式を得る．

$$(\mathbf{d}_t, \mathbf{s}'_t) = f_{temporal}(\mathbf{d}_{t-1}, \mathbf{s}'_{t-1}, \tilde{y}_{t-1}) \tag{9}$$

最後に，長さ T の多変量説明変数時系列 $X_T = (\mathbf{x}_1, \cdots, \mathbf{x}_t, \cdots, \mathbf{x}_T)$（ただし $\mathbf{x}_t = (x_t^1, \cdots, x_t^n)$）および目的変数時系列 $\mathbf{y}_T = (y_1, \cdots, y_t, \cdots, y_T)$ から時間 $T+1$ の目的変数 y_{T+1} を予測するため，隠れ状態 $\mathbf{d}_T$ とコンテキストベクトル $\mathbf{c}_T$ を連接し，次式を用いる．

$$\hat{y}_{T+1} = F(X_T, \hat{\mathbf{y}}_T) \tag{10}$$

$$= \mathbf{v}_y^{\mathrm{T}}(W_y[\mathbf{d}_T; \mathbf{c}_T] + \mathbf{b}_y) + b'_y \tag{11}$$

ここで，F は予測器を示す．$W_y \in \mathbb{R}^{p \times (p+m)}$ および $\mathbf{b}_y \in \mathbb{R}^p$ は連接ベクトル $[\mathbf{h}_T; \mathbf{c}_T] \in \mathbb{R}^{p+m}$ を p 次元潜在空間に写像している．重み $\mathbf{v}_y \in \mathbb{R}^p$ およびバイアス $b'_y \in \mathbb{R}$ をともなう線形変換により，最終的な目的変数の予測値 $\hat{y}_{T+1}$ を求めている．

3 データ

3.1 財務時系列データ

本論文では，統計法に基づく基幹統計として「法人企業統計調査規則」（昭和 45 年大蔵省令 48 号）に基づいて財務省により収集された「法人企業統計調査」を使用する[2]．2003 年の第 1 四半期から 2016 年の第 4 四半期までの 14 年間（56 四半期）の調査結果を使用する．なお，本調査は年次調査と四半期調査で構成されているが，四半期調査を主に使用し，年次調査は補完処理のために補助的に用いる．本調査に収録された企業の総数は四半期調査で 57,775，年次調査で 60,516 であった．調査項目を表 1 に示す．

四半期調査による財務諸表に基づいて求めた種々の財務比率を目的変数または説明変数として用いる．前処理として，まず，短期間のみに記録された企業などのデータを除外する（除外処理）．次に，財務比率を計算する前に財務諸表における欠損値を補完する処理を行う（補完処理）．最後に，補完済みの財務諸表を用いて財務比率の計算を行う（計算処理）．以上の手順によって得られた四半期次の各種財務比率の時系列を分析に用いる．それぞれの詳細について以下に示す．

除外処理　次のいずれかの条件に該当する企業データを除外した．すなわち，除外の対象は，(1) 財務諸表の分析対象期間（56 期）の一部において全調査項目が欠損している企業データ，(2) 財務諸表におけるある調査項目に関して分析対象の全期間にわたって欠損している企業データである．この除外処理の後，分析に使用される企業数は 2296 となった．なお，後述する業種特化型モデルのために着目する卸売業の企業数は 237 となった．

補完処理　本論文における補完処理は，年次データを利用した補完と線形補間に基づく．詳細については稿を改めて報告する．

計算処理　本論文では種々の財務比率を説明変数または目的変数として使用する．各財務比率は表 2 で定義される通りであり，これらは基本的に「法人企業統計調査」において使用されているものである．

3.2 マクロ経済時系列データ

本論文では，マクロ経済時系列データとして日経平均株価終値（以下，日経平均）および日本の GDP を用いる．ただし，日経平均は 2003 年 1 月から 2016 年 12 月までの月次データ[3]を用いる．また，日本の GDP は 2003 年第 1 四半期から 2016 年第 4 四半期までの四半期次データ[4]を用いる．

4 提案手法

本章で述べる提案モデルは 3 つの特徴を持つ．1 つ目は多変量時系列が複数与えられる状況すなわち企業ごとに複数の財務比率で表現された時系列が企業数だけ与えられる状況を仮定したモデルである点であり，2 つ目は財務時系列に加えてマクロ経済時系列を反映したモデルである点，3 つ目は対象業種の動向を反映した業種特化型モデルである点である．また，提案モデルは目的変数時系列の 1 期先を予測するようなモデルである．提案モデルにおける入力された多変量時系列から予測出力に至るまでの流れを図 4 に示す．第 1 フェー

[2]調査対象のうち，「金融および保険」業界の企業は除外した．

[3]https://indexes.nikkei.co.jp/nkave/index/profile?idx=nk225

[4]https://www.esri.cao.go.jp/jp/sna/data/data_list/sokuhou/files/2019/toukei_2019.html

表 1: 財務諸表．ただし，「†」は指定項目が期首に記録されることを示し，「‡」は期末に記録されることを示す．

分類	四半期調査	年次調査
資産	支払手形 †,‡	支払手形 †,‡
	買掛金 †,‡	買掛金 †,‡
固定資産	土地 †,‡	土地 †,‡
	建設仮勘定 †,‡	建設仮勘定 †,‡
	その他有形固定資産 †,‡	その他有形固定資産 †,‡
	無形固定資産 †,‡	無形固定資産（ソフトウェア以外） †,‡
		ソフトウェア †,‡
	総資本 †,‡	総資本 †,‡
人員	従業員数 ‡	従業員数 ‡
損益	減価償却費 ‡	減価償却費 ‡
		特別減価償却費 ‡
	売上 ‡	売上 ‡
	売上原価 ‡	売上原価 ‡
	営業利益 ‡	営業利益 ‡
	経常利益 ‡	経常利益 ‡

表 2: 財務比率．ただし，「$*$」は指定項目について各四半期の期首と期末の平均（すなわち期中平均）が用いられることを示し，「$**$」は指定項目について各四半期の期首から期末への増減量が用いられることを示す．

記号	財務比率	計算式
X_0	総資本営業利益率	$\frac{\text{営業利益}}{\text{総資本}^*}$
X_1	総資本経常利益率 (ROA)	$\frac{\text{経常利益}}{\text{総資本}^*}$
X_2	売上高営業利益率	$\frac{\text{営業利益}}{\text{売上}}$
X_3	売上高経常利益率	$\frac{\text{経常利益}}{\text{売上}}$
X_4	総資本回転率	$\frac{\text{売上}}{\text{総資本}^*}$
X_5	有形固定資産回転率	$\frac{\text{売上}}{(\text{土地}+\text{その他有形固定資産})^*}$
X_6	買掛金回転率	$\frac{\text{売上}}{\text{支払手形}+\text{買掛金}}$
X_7	減価償却率	$\frac{\text{減価償却費}}{\text{その他有形固定資産}+\text{有形固定資産}+\text{減価償却費}}$
X_8	労働装備率	$\frac{(\text{土地}+\text{その他有形固定資産})^*}{\text{従業員数}}$
X_9	キャッシュフロー率	$\frac{\text{経常利益}+\text{減価償却費}}{\text{総資本}^*}$
X_{10}	設備投資率	$\frac{\text{建設仮勘定}^{**}+\text{その他有形固定資産}^{**}+\text{無形固定資産}^{**}+\text{減価償却費}}{\text{総資本}^*}$
X_{11}	粗利率	$\frac{\text{売上}-\text{売上原価}}{\text{売上}}$

ズは対象業種の多変量時系列における空間的相関を抽出することを目的とする．第 2 フェーズは説明変数時系列と既知の目的変数時系列における空間的相関を抽出することを目的とする．第 3 フェーズは時間的関連性を抽出し，時間 $1, \cdots, T$ の多変量時系列が与えられたとき，時間 $T+1$ の目的変数を予測することを目的とする．以下に各フェーズについて詳述する．

第 1 フェーズ　このフェーズでは，長さ L の多変量時系列における空間的相関を抽出する．多変量時系列データから固定幅ウィンドウの部分時系列を切り出す前に，その時系列データ全体に対してアテンションメカニズムを適用することによって，多変量時系列間の空間的相関をより適切に抽出する点に狙いがある．多変量時系列が複数与えられた状況では，固定幅の部分時系列を切り出した後にアテンションメカニズムを適用するならば，同一企業の複数の部分時系列と他企業の部分時系列が同列に扱われて学習されるため，企業という単位を考慮した空間的相関を抽出するのに適切でないと考えられるためである．そこでまず企業を単位とした複数の多変量時系列に対して空間的アテンション LSTM を適用し，時系列間の空間的相関を抽出する．J を企業数，L を学習データの時系列長，K を多変量時系列の変量数とすると，多変量時系列 A は $J \times L \times K$ のテンソル（以下，(J, L, K) と表記する）で表現できる．これに空間的アテンション LSTM を適用する．これに対して式 (1)，(2)，(3) と同様に空間的アテンション重みによる重みづけ（以下，$F_{spatial}$ と表記）を適用する．

$$\acute{A} = F_{spatial}(A) \tag{12}$$

その後，W を重み，b をバイアスとする線形変換により $(1, L, K)$ の $\tilde{A}$ へと集約する．

$$\tilde{A} = W\acute{A} + b \tag{13}$$

また，日経平均 N は月次データを用いるとき時系列長は $3L$ であるが，アテンションメカニズムを適用するとともに四半期次に集約する．このとき，$(1, 3L, 1)$ の N から $(1, L, 1)$ の $\tilde{N}$ へと集約される．具体的には各四半期 t に対して 1 月めから 3 月めまでを $i \in \{1, 2, 3\}$ で

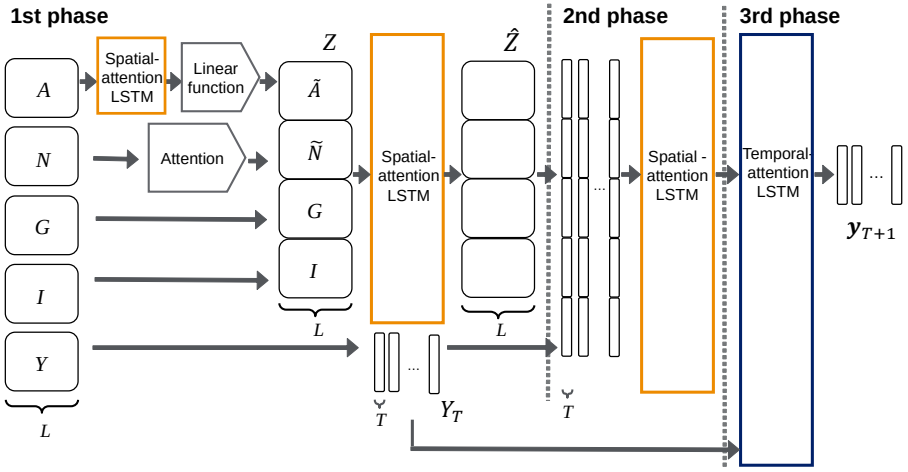


図 4: 提案モデルの構造.

示すとき次式を用いる．

$$\gamma_t^i = \frac{\exp(N_t^i)}{\sum_{j=1}^{3}\exp(N_t^j)} \tag{14}$$

$$\tilde{N}_t = (\gamma_t^1 N_t^1, \gamma_t^2 N_t^2, \gamma_t^3 N_t^3)^{\mathrm{T}} \tag{15}$$

以上のようにして全企業の財務時系列が集約された $\tilde{A}$ に対して，四半期次に集約された日経平均の時系列 $\tilde{N}$，日本の GDP の時系列 G，および，対象業種における説明変数時系列 I（目的変数時系列を含まない）を連接する．なお，G は四半期次データであるため，そのサイズは $(1, L, 1)$ である．

$$Z = [\tilde{A}; \tilde{N}; G; I] \tag{16}$$

ここで，学習時のミニバッチ数を B とすると，I のサイズは $(B, L, K-1)$ である．したがって，連接後の Z のサイズは $(B, L, 2K+1)$ となる．これに対して空間的アテンション重みで重みづけることにより，時系列間の空間的相関を抽出する．

$$\hat{Z} = F_{spatial}(Z) \tag{17}$$

なお，式 (12) と (17) の $F_{spatial}$ は互いに独立に，それぞれ A と Z から求める．

その後，長さ L の時系列 $\hat{Z}$ の開始期から固定長 T の部分時系列を切り出し，1 期ずつずらして固定長 T の部分時系列の切り出しを順次繰り返し，それらを次のフェーズの入力とする．したがって，次のフェーズの入力サンプル数は $B \times (L-T)$ となり，そのサイズは $(B(L-T), T, 2K+1)$ となる．

第 2 フェーズ　第 2 フェーズでは，まず，第 1 フェーズで得た長さ T の $\hat{Z}$ と目的変数時系列 Y_T を，各期 $t \in \{1, \cdots, T\}$ で対応づけて連接する．ここで，第 1 フェーズの最後に述べた $\hat{Z}$ を長さ T で分割する要領で，Y_T も長さ T で分割した部分時系列を用いる．

$$\dot{Z} = [\hat{Z}; Y_T] \tag{18}$$

ここで，連接後の $\dot{Z}$ のサイズは $(B(L-T), T, 2K+2)$ となる．これに対して空間的アテンション LSTM を適用し，目的変数時系列と説明変数時系列の間の空間的相関を抽出する．

$$\hat{\dot{Z}} = F_{spatial}(\dot{Z}) \tag{19}$$

第 3 フェーズ　このフェーズでは，時間的アテンション LSTM を適用し，第 2 フェーズで得た空間的相関の時間的関連性を捉える．つまり，異なる時間の異なる（部分）時系列の間の時空間関連性を抽出し，式 (10) と同様に時間的アテンション LSTM に基づく予測器（以下，$F_{temporal}$ と表記）で 1 期先の目的変数を予測する．

$$\hat{\mathbf{y}}_{T+1} = F_{temporal}(\hat{\dot{Z}}, Y_T) \tag{20}$$

この $\hat{\mathbf{y}}_{T+1}$ が最終的な予測結果である．$\hat{\mathbf{y}}_{T+1}$ は予測対象の企業数だけの長さを持つベクトルであり，時間 $T+1$ での各企業に対する目的変数予測値を要素にとる．

提案モデルの学習には誤差逆伝播アルゴリズムを用いる．学習過程で Adam オプティマイザ [7] によって，予測ベクトル $\hat{\mathbf{y}}_{T+1}$ と正解ベクトル $\mathbf{y}_{T+1}$ の間の平均二乗誤差（MSE）を最小化する．

モデル構造　第 1 フェーズにおいて，部分時系列の切り出しを行う前に空間的相関を抽出し，企業を単位とした相関関係を捉えるようなモデル構造を設計した．その上で，前述の第 1〜第 3 フェーズの機能を実現させ

て，多変量説明変数による時系列が複数与えられた状況での予測をより適切に行うことを狙いとした．それに対して，Liu ら [10] の DSTP モデルの研究においては，多変量説明変数による時系列が一つ与えられた状況が想定されている．この DSTP モデルを本論文で想定する多変量説明変数による時系列が複数与えられた状況下で用いる場合，同一企業の複数の部分時系列と他企業の部分時系列が同列に扱われて学習されることになる．提案モデルは，企業を単位とした時系列によって学習される点が特徴的であると言える．

5 実験

5.1 実験設定

第 4 章で述べた提案モデルの有効性を評価するため，従来の DSTP モデル [10] および単純な LSTM モデルと比較する．本実験では対象業種として「卸売業」（Whole sale and trade）を仮定する．また，予測対象の目的変数を総資本経常利益率（ROA）すなわち $Y = X_1$ とし，説明変数を $X = \{X_0, X_2, X_3, \cdots, X_{11}\}$ とする．本実験は検証実験とテスト実験からなる．いずれの場合も，実験に用いるデータセットの時系列において開始期から固定幅ウィンドウの部分時系列を抽出し，次に 1 期先にスライドさせて同じく固定幅ウィンドウの部分時系列を抽出することを順次繰り返すことにより，合計 15 の部分時系列を作成する．

検証実験　40 期からなる部分時系列セット ($L = 40$) によりモデルを学習し，それぞれの 1 期先を予測する際の誤差に基づいて，モデルのハイパーパラメータを最適化する．このとき，ハイパーパラメータとして特に重要な中間ユニット数（隠れ状態数）U に着目する．その探索範囲は $U \in \{16, 32, 64, 128\}$ とする．

テスト実験　検証実験で選択したハイパーパラメータを用いて，41 期からなる部分時系列セット（$L = 41$）によりモデルを学習し，それぞれの 1 期先を予測する際の誤差の平均に基づいて予測性能を評価する．

その他の実験設定と評価指標　中間ユニット数 U 以外のハイパーパラメータについては，時間ウィンドウの長さを $T = 12$，エポック数を 500，学習率を 0.001，バッチサイズを 64 とした．また，単純な LSTM モデルにおいて中間層の数は 2 とした．これらは検証実験の過程で経験的に選択した．

従来の DSTP モデルと単純な LSTM モデルでは，すべての入力時系列が四半期次である必要があるため，マクロ経済時系列として用いる日経平均の月次データについて 3ヶ月分の平均をとることで四半期次とした．

本実験における評価指標には平均二乗誤差（mean squared error: MSE）を用いた．

5.2 結果と考察

表 3 に，各種設定で提案モデル，DSTP モデル，および単純な LSTM モデルを用いた平均二乗誤差（MSE）および標本標準偏差（SD）による評価結果を示す．表中の「Data used」には各モデルに用いた説明変数時系列の種類を示す．なお，これらに加えてすべてのモデルに（既知の）目的変数時系列を用いている．全業種の財務時系列 A と対象業種の財務時系列 I は，$I \subset A$ の関係にあることに注意する．また,「# of units」には検証実験を経て選択された各モデルの中間ユニット数を示す．以下では，表 3 の実験結果に基づいて 3 つの観点から考察と貢献を述べる．なお，比較評価の際はデータ量に関して公平を期している．評価の結果，以下に示す比較のすべてにおいて，ウィルコクソンの符号付き順位検定による統計的な有意差がみられた（有意水準 0.05）．

マクロ経済時系列を反映したモデル構造による効果　マクロ経済時系列として日経平均（N）と日本の GDP（G）を反映した Model 1 はそうでない Model 3 を上回っている．これは提案モデルが適切にマクロ経済時系列を活用できていることを示唆する．

業種に特化した構造による効果　まず初めに，単純 LSTM モデルにおいて，対象業種の財務時系列を用いた Model 9 の予測性能が，全業種の財務時系列を用いた Model 8 のそれを上回った．これは LSTM において業種の動向に着目することが有効であるという仮説の妥当性を示唆する．一方，DSTP モデルにおいては，対象業種の財務時系列を用いた Model 5 は全業種の財務時系列を用いた Model 4 を下回った．これは対象業種の財務時系列に限定した場合に学習のためのデータ量が十分でなくなり，そのことが単純 LSTM よりも複雑な構造をそなえた DSTP モデルにおいて顕在化したであろうことが理由として考えられる．また，全業種の財務時系列 A を含めたすべての時系列データを用いた提案モデルの Model 1 は，A によらない Model 2 を上回った．これは提案モデルにおいて式 (12) および (13) による全業種の集約が，すべての時系列データの特徴をモデルに組み込む上で有効であることを示唆する．

多変量説明変数による時系列が複数与えられた状況を考慮してモデル化することの効果　最後に，すべての時系列データを用いた提案モデルの Model 1 は DSTP モデルの Model 4 および単純 LSTM モデルの Model 7

表 3: 種々の条件下での提案モデルと比較対象モデルによる平均二乗誤差（MSE）と標本標準偏差（SD）.

Model IDs	Models	Data used	# of units	MSE	SD
Model 1	Proposed model	(I), A, N, G	16	2.08×10^{-4}	8.48×10^{-5}
Model 2	Proposed model (w/o 'A')	I, N, G	32	2.11×10^{-4}	7.43×10^{-5}
Model 3	Proposed model (w/o 'N' and 'G')	(I), A	128	2.28×10^{-4}	7.64×10^{-5}
Model 4	DSTP-based model	A, N, G	16	2.12×10^{-4}	6.47×10^{-5}
Model 5	DSTP-based model	I, N, G	16	2.26×10^{-4}	9.40×10^{-5}
Model 6	DSTP-based model	A	16	2.11×10^{-4}	7.43×10^{-5}
Model 7	Simple LSTM model	A, N, G	16	3.86×10^{-4}	6.07×10^{-4}
Model 8	Simple LSTM model	A	16	3.10×10^{-4}	3.72×10^{-4}
Model 9	Simple LSTM model	I	16	2.95×10^{-4}	3.43×10^{-4}

を上回った．その一方で，マクロ経済時系列を用いない状況下では，提案モデルである Model 3 が DSTP モデルである Model 6 の性能を下回る結果となった．多変量説明変数からなる時系列が複数与えられる状況すなわち企業ごとに複数の財務比率で表現された時系列が企業数だけ与えられる状況において，適切なモデル構造についてさらなる検討の余地がある．

6 むすび

企業財務時系列を包括的に分析・予測するための有用な手段を確立するため，本論文では多変量説明変数で表現された時系列が複数与えられた状況を仮定したモデルを設計し，目的変数の 1 期先の予測を行う問題を想定した検討を行った．この目的の下で，マクロ経済時系列を適切に組み込んだ LSTM に基づく新たなモデルを提案した．また，対象業種に固有の動向を反映しつつ，全業種の企業に関する財務時系列の特徴を組みこむことができるモデル構造を設計した．提案モデルの有効性に関して，次に示す 3 つの貢献が得られた．一つ目に，提案モデルは GDP などのマクロ経済時系列を組み込みつつ対象業種の動向を反映した場合において，従来の DSTP モデルおよび LSTM モデルよりも多変量説明変数からなる複数時系列から目的変数時系列の 1 期先を適切に予測できることを実験的に示した．しかしながら，財務時系列のみを用いて予測を行った場合においては，DSTP モデルの方がより適切に予測できたことから，モデル構造にさらなる改善の余地があると言える．二つ目に，提案モデルにおいてはマクロ経済時系列を適切に反映して予測に活用することができた．三つ目に，提案モデルでは，対象業種に固有の動向を反映しつつ，全業種の企業財務時系列の特徴を組みこむことができた．なお，本論文の実験では対象業種として卸売業を取りあげ，目的変数時系列として ROA に着目したが，他の業種や他の目的変数についての検討は今後の課題である．

謝辞 本研究の一部は科学研究費補助金基盤研究 (B)（15H02703）の援助による．

参考文献

[1] Amini, M. H., Kargarian, A. and Karabasoglu, O.: ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation, *Electric Power Systems Research*, Vol. 140 (2016), 378–390.

[2] Choi, E., Bahadori, M. T., Schuetz, A., Stewart, W. F. and Sun, J.: RETAIN: Interpretable predictive model in healthcare using reverse time attention mechanism, *Advances in Neural Information Processing Systems*, Vol. 29 (2016), 3504–3512.

[3] Goodfellow, I., Bengio, Y. and Courville, A.: *Deep Learning*, MIT Press, 2016.

[4] Hochreiter, S.: Untersuchungen zu dynamischen neuronalen Netzen, Diploma thesis, Institut für Informatik, Technische Universiät München, 1991.

[5] Hochreiter, S., Bengio, Y., Frasconi, P. and Schmidhuber, J.: Gradient flow in recurrent nets: the difficulty of learning long-term dependencies, in Kremer, S. C. and Kolen, J. F. eds., *A Field Guide to Dynamical Recurrent Neural Networks*, IEEE Press, 2001.

[6] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, *Neural Computation*, Vol. 9 (1997), 1735–1780.

[7] Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, in *Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)*, 2015.

[8] Li, H., Shen, Y. and Zhu, Y.: Stock Price Prediction Using Attention-based Multi-Input LSTM, *Proceedings of Machine Learning Research*, Vol. 95 (2018), 454–469.

[9] Liang, Y., Ke, S., Zhang, J., Yi, X. and Zheng, Y.: GeoMAN: Multi-level attention networks for geo-sensory time series prediction, in *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI-18)*, 2018.

[10] Liu, Y., Gong, C., Yang, L., Chen, Y.:DSTP-RNN: A dual-stage two-phase attention-based recurrent neural network for long-term and multivariate time series prediction, *Expert Systems with Applications*, Vol. 143 (2019).

[11] Qin, Y., Song, D., Cheng, H., Cheng, W., Jiang, G. and Cottrell, G. W.: A dual-stage attention-based recurrent neural network for time series prediction, in *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI-17)*, 2017.

経済金融時系列における歪みの検出

Detection of distortion in economic and financial time series

土屋 太助 [1]

Taisuke Tsuchiya[1]

Abstract: The market is greatly disrupted by economic news and online media information such as twitter, causing distortions in the economic and financial time series of commodity futures markets. In this research, we applied the anomaly detection method by using of Dense Auto Encoder to economic and financial time series data to detect distortion of economic and financial time series such as commodity futures market.

1. はじめに

近年、金融データへの AI の応用が注目されており、経済金融時系列を Recurrent Neural Network（RNN）の入力として経済金融時系列の期先予測を行うといった研究[1]が行われているが、経済ニュースや twitter といったオンラインメディアの情報により、市場は大きく混乱し、各国の経済政策や地政学上の問題によってはこれまでの価格決定ロジックの常識が覆ることがある。また、オンラインメディアの情報を入力として自然言語処理により経済金融時系列の期先予測を行う研究[2]も行われているが、オンラインメディアの情報のすべてが常にターゲットの時系列に直接的に影響を与えるわけではなく、例えば、「NY ダウが暴落し金が高騰」といった、金融時系列からも得られる重複した二次情報も含まれている。金融時系列とオンラインメディアから抽出したスパースな特徴量を単純に結合しても、期先予測の精度向上は容易ではない. その時々の相場感、経済環境、ニュースを鑑みて、専門家が特徴量を都度変更し、モデルを改修して運用することも可能だが、メンテナンス性、汎用性の観点から実用的ではない。加えて、今現在予測精度が高くても、そのモデルがいつまで、どういう状態であれば使い続けても問題ないのか、どういう状態になったら使用を停止すべきかの判断も難しい。データからこうした状況変化のシグナルを検出できれば、（部分的にでも）期先予測モデルの運用開始を意思決定することができるのではないかと考えた.

本研究は、経済金融時系列から目的変数毎のレジームの転換点を検出し、レジーム毎に最適なモデル（モデルの使用停止を含む）を選択することを最終目的としている. 今回はその中で、異常検出手法を用いた経済金融時系列の歪みの検出を行った.

具体的には、センサデータの異常検知等[3]で用いられる自己符号化器（Auto Encoder）を用いて経済金融時系列の歪み（異常）を検出する. 自己符号化器により次元圧縮し、入力と出力の差分である再構築誤差を金融時系列の歪みと定義し解析した.

次に、歪み情報を予測モデルの入力データと結合することで、予測精度改善効果が得られるか否かを検証した.

2. 自己符号化器による異常検知

オートエンコーダはニューラルネットワークの一種で、情報量を小さくした特徴表現を獲得する. オートエンコーダによる異常検知の仕組みは、エンコーダで入力(x)を、入力(x)より低次元な潜在変数(z)に次元圧縮し、デコーダで元の入力を復元する. 学習は教師なし学習で、入力と出力の誤差が小さくなるように学習を行う. 異常検知はこの入力と出力の誤差（再構築誤差）を用いる. (図 1)

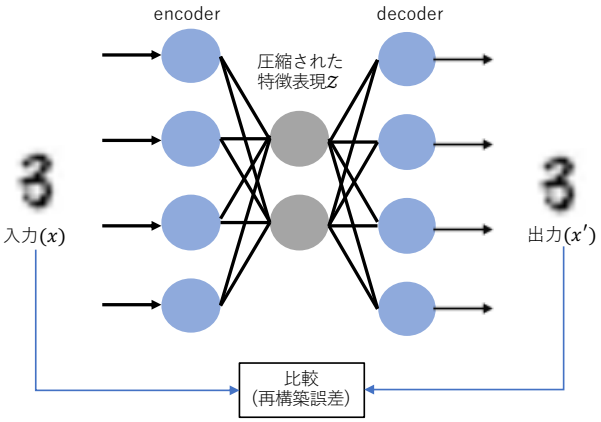


図 1：自己符号化器による異常検知の仕組み

推論時、学習データに含まれない未知のデータ入力に対して再構築誤差が大きいときに異常として検出する仕組みである．センサデータや画像データの異常検知に用いられる手法だが、画像系の異常検知においては全結合層ではなく、畳み込み層（Convolution 層）を、Deep なネットワーク構造とともに用いることが一般的である．Resnet や skip-connect を追加した U-net など復元能力の高いネットワーク構造は、未知のデータに対しても再構築能力が高いため、再構築残差が出にくく、結果として異常を検出できない。データおよび異常の特性を踏まえ、適度な復元力のネットワーク構造、ハイパーパラメータを選ぶ必要がある．

3. モデルへの入力特徴量

ニューヨーク商品取引所（NYMEX）の原油先物価格をターゲットとし、金、プラチナ、パラジウムなどの先物価格に加えて日本、米国、中国、欧州の為替、米国、中国の主要な株価指数の計 38 種の日次データを特徴量として使用した．

4. ML アーキテクチャと具体的な実装

アーキテクチャ全体はデータ取得、前処理、予測の 3 つプロセスで構成しており、Dense Auto Encoder による次元圧縮および再構築残差の算出し、オートエンコーダのニューラルネットワークの構造は入力層、隠れ層 1 層、出力層という深層ではないネットワーク構造を用いた．歪みの検出という目的を鑑みて、適度な残差が発生するよう、全結合の隠れ層 1 層のシンプルなネットワーク構造を選択する．

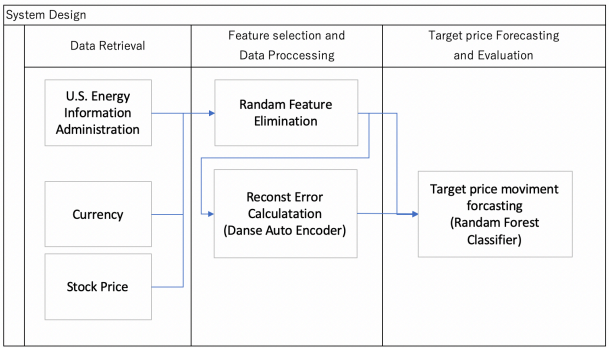


図 2：ML アーキテクチャ概要

具体的なモデルとして下記を用いる．

1. Dense Auto Encoder: TensorFlow を用いて実装を行った．隠れ層 1 層の全結合型の構造で、活性化関数は Tanh 関数を用いた．epoch 数は 200epoch とし、最 適化アルゴリズムに Adam を用い、loss 関数は MSE を用いた．
2. Random Feature Elimination(RFE): scikit-learn の sklearn.feature_selection を用いて実装を行った．
3. Random Forest (RF):パッケージ scikit-learn のクラス Random Forest Classifier を用いて実装を行った．

5. 結果と考察

まず、ターゲットである原油先物価格のテスト期間における再構築誤差の推移を図 3 にて示す．

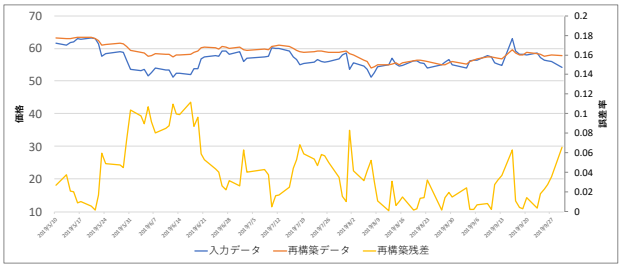


図 3：テスト期間における再構築誤差の推移

Autoencoder の入力 38 次元に対して、8 次元まで次元圧縮しているため、テスト期間の時間が進むにつれて再構築誤差が発散する可能性もあると予想していたが、予想に反してテスト期間のターゲット先物価格の再構築誤差は一定以下に抑えられており、次元圧縮、特徴抽出は成功していると言える．
また、2019 年 5 月下旬から再構築誤差が 1 ヶ月程度大きくなっているが、これらの時期に、トランプ大統領による原油の増産圧力をかける発言がでており、実際に入力データに含まれていない米国在庫統計の情報を確認したところ供給量増（価格下落要因）が確認できた。

次に、入力に在庫統計情報を加え、1 日先のターゲット価格の価格上昇、下落を Random Forest Classifier により 2 クラス分類問題として予測した．入力データに Dense Auto Encoder の次元圧縮後特徴量、再構築誤差を結合させ Random Forest Classifier モデルに入力した。評価結果(30 回評価の平均値)を図 4 に示す．入力＋再構築誤差（ターゲットの誤差のみ）を入力としたモデルがベースモデルよりもわずかに精度が高いという結果が出た．一方で全特徴量分の再構築誤差を入力に加えた場合、精度は向上しなかった．

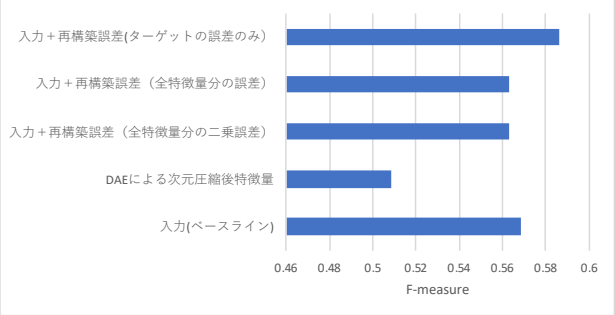


図 4：DAE 特徴量を加えた際の F 値変化

6. まとめと今後の課題

本研究では、経済金融時系列データから歪みを検出し、レジーム毎にモデルを構築し、最適なモデルを自動選択することを最終目的として、今回はその中で、自己符号化器を用いた経済金融時系列からターゲット変数毎の歪みを抽出できることを確認した．今後の課題を下記に列挙する．

・自己符号化器を用いた Out of Distribution の検出
・再構築誤差を用いた解釈可能性の高いレジーム分解手法の構築
・レジーム毎の動的特徴選択とモデル構築

また、今回は、少ない入力（特徴次元数）で実験したが、Autoencoder による次元圧縮の効果を最大化するためには精度向上に寄与する意味のある特徴量の追加が必要である．

参考文献

[1] Wei Bao, Jun Yue, Yulei Rao, A deep learning framework for financial time series using stacked autoencoders and long-short term memory(2017)
[2] Xuerong Li, Wei Shang, Shouyang Wang : Text-based crude oil price forecasting: A deep learning approach(2018)
[3] Chong Zhou, Randy C. Paffenroth : Anomaly Detection with Robust Deep Autoencoders, KDD '17 the 23rd ACM(2017)

ニュース重要単語の機械学習によるアクティブ運用

Active Investment Management by Machine Learning of Important News Keywords

三好勝博[1*]　細木唯以[1]　江口潤一[2]　鈴木智也[1,2]
Masahiro Miyoshi[1] Yui Hosoki[1] Junichi Eguchi[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] 大和証券投資信託委託株式会社
[2] Daiwa Asset Management Co.Ltd.

Abstract: In stock investment and asset management, numerical information has been mainly used for fundamentals analysis and technical analysis, but recently text information such as news released in real time can be also used by development of natural language processing techniques. Although machine learning is mainly used for investment decisions, because text information is composed of a large number of words, its bag-of-words tend to be a sparse and high-dimensional vector. Therefore, the curse of dimensionality causes a negative effect on the learning performance of machine learning. For this reason, we only focus on news headlines to restrict the number of keywords as text information. In addition, we extracted important keywords that increase the volatility of stock prices right after the news appears, and applied machine learning techniques to learn the relationship between the combination of important keywords included in a news headline and tomorrow's active return of the company most related to the news. Through some statistical tests, we could confirm the validity of focusing on the stock volatility to extract important keywords and their combination is useful for active investment management with machine learning approach.

1　はじめに

近年では AI による自然言語処理技術が発展し，イベント的に発生するニュースなどのテキスト情報もアルゴリズムに組み込んだ投資判断が活発になりつつある [1, 2, 3]．投資判断には主に機械学習を用いるが，テキスト情報は多数のワードで構成されているため Bag-of-words の局所表現によりベクトル化すると高次元でスパースになりやすい．その結果，次元の呪いにより機械学習の学習性能に悪影響を及ぼす．

そこで本研究では，キーワードの種類を抑制するため，テキスト情報の中から将来の株価変動 (ボラティリティ) に大きく影響する重要キーワードを抽出する．その後，抽出した重要キーワードを用いて，各ニュースヘッドラインに含まれる重要キーワードの組み合わせと翌日のアクティブリターンを機械学習し，アクティブ運用による投資シミュレーションを行う．なおアクティブ運用とは，ベンチマーク (本研究では TOPIX) を上回る運用パフォーマンスを目指す運用手法である．

*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: 18nm500x@vc.ibaraki.ac.jp

2　ニューステキストの前処理

本研究では使用するテキスト情報として，QUICK ニュース解析サービス [4] から配信されるニュースヘッドラインを用いる．形態素解析ツールである MeCab[5] を用いてニュースヘッドラインを形態素に分割し，重要キーワードの候補を得る．ヘッドラインはニュース本文と比べテキスト量が少ないが，重要な情報およびキーワードが凝縮されている．さらにテキスト情報をヘッドラインに限定することで，登場するキーワードを抑制する．

ヘッドライン中には●や★などの装飾記号や数字が混在し，ワード数の増加を招くと考えられる．そのため，予めこれらのワードを削除する．また本研究では，個別銘柄に対してアクティブ運用を行うため，ニュースの配信元が「TECH」「QEC」「日銀」「財務省」であるヘッドラインや，銘柄コードに「日銀」「JPX」「日証金」を含むヘッドラインは対象外とする．

ヘッドライン中には共起関係の高いキーワード (共起ワード) が存在する．共起関係とは，ある単語がある文章中に出現した際に，その文章中に特定の単語が頻出することを指す．例えばエアリズムなど，特定の銘柄 c (この場合はユニクロ) のみに対応するキーワードは，株価変動との関係を汎用的に評価できない．そこで，式 (1) を満たすキーワード k を共起ワードとみなし，予めこれらも削除する．

$$P(c|k) = \frac{N(c|k)}{N(k)} \geq 0.5 \tag{1}$$

ここで $N(k)$ はキーワード k を含むニュース数，$N(c|k)$ はキーワード k を含むニュースが銘柄 c と最も関連した回数である．つまり式 (1) は，キーワード k を含むニュースの半分以上が銘柄 c のものであることを意味する．なお，銘柄コードが付与されていないニュースは c が不明なので対象外とする．

3　重要キーワードの抽出

3.1　ボラティリティ変化率

本研究では，株価ボラティリティに注目し，重要キーワードを抽出する．金融市場に大きな影響を与えるニュースが出現すると，そのニュースに関連する株価は大きく変動することが予想される．そこで本研究では，ニュース出現前後の株価ボラティリティの変化を計算し，ニュースヘッドラインを構成するキーワードの重要性を評価する．このボラティリティ変化が高いほど，重要なキーワードとみなす．

まず，出現したニュースと関連がある企業の株価を $x(d)$ とする．本研究では日次終値を用い，d は日付を意味する．株価の収益率 $r(d)$ は

$$r(d) = \frac{x(d) - x(d-1)}{x(d-1)} \tag{2}$$

であり，直近 D 日間の標準偏差を計算することで，ボラティリティ $v(d)$ を得る．

$$v(d) = \sqrt{\frac{1}{D}\sum_{i=0}^{D-1}\{r(d-i) - m(d)\}^2} \tag{3}$$

ここで $m(d)$ は直近 D 日間の収益率の平均値である．なお本研究では，1 カ月の営業日として $D = 20$ と設定した．

ニュース発生前後のボラティリティ変化を評価する際に，他媒体で既にニュースが報道された可能性を考慮するため，時間を遡るパラメータ b を導入する．ニュース発生前のボラティリティを $v(d-b)$ とすると，ニュース発生前後のボラティリティ変化 $V(d)$ は，

$$V(d) = \frac{v(d)}{v(d-b)} \tag{4}$$

として評価される．発生したニュースを構成する全キーワードに対して，この変化率 $V(d)$ を紐づける．

実際のニュースヘッドラインはリアルタイムに配信されるが，ニュースが配信されるたびに投資判断および売買注文を自動化するには，高額な設備投資が必要である．そこで本研究では試作的な運用として，毎日 12 時に直近 24 時間分のニュースデータを一括で取得し，得られたニュースに基づいて当日終値付近で売買を執行する．したがって，当日 d 日で得られるニュースには，前日 $d-1$ 日のニュースの一部が含まれている．よって，式 (4) に示すボラティリティ変化率 $V(d)$ の計算において $b \geq 2$ にする必要がある．他媒体での先行発表も考慮すると b を延長する必要があるが，本研究では最も基礎的な設定として $b = 2$ とする．

3.2　重要キーワードの検定

時間経過とともに様々ニュースが発生するため，各キーワード k $(k = 1, 2, \cdots, K)$ において変化率 V に関する集合 $\mathbf{G}_k$ を得る．この真の母集団は未知であるが，母平均を μ_k，母分散を σ_k^2 と書く．

まず，集合 $\mathbf{G}_k$ の平均値 $\bar{V}_k$ は母平均に関する不偏推定量となるため，$\bar{V}_k$ をキーワード k におけるボラティリティ変化率と考えることができる．しかし各キーワードにおいて出現回数 n_k は異なり，n_k が小さいほど推定値の信頼性は低下する．つまり $\bar{V}_k$ の大小を単純に比較することで，重要キーワードを選出できるとは言い難い．

そこで，母平均の検定を利用する．中心極限定理より，平均値 $\bar{V}_k$ は $N\left(\mu_k, \sqrt{\frac{\sigma_k^2}{n_k}}\right)$ に従う．キーワードとして重要性が無く，ボラティリティ変化が無い $(\mu_k = 1)$ ことを帰無仮説 H_0 とすると，帰無仮説 H_0 と集合 $\mathbf{G}_k$ のマハラノビス距離によって，検定統計量である t 値を得る．

$$t_k = \frac{\mu_k - 1}{\sqrt{\frac{\sigma_k^2}{n_k}}} \tag{5}$$

ただし μ_k と σ_k^2 は未知であるため，それぞれ平均値 $\bar{V}_k$ および不偏分散 $s_k^2 = \frac{1}{n_k - 1}\sum_{j=1}^{n_k}(V_{k,j} - \bar{V}_k)^2$ を用いる．この t_k に基づいて大小比較すれば，サンプルサイズ n_k の違いを考慮できる．本研究では有意水準 5% (片側検定) を満たすキーワード k に厳選する．

ただし n_k が少なすぎるキーワードについては，先に除去しておくことが賢明である．そこで本研究では，

10 日間に 1 回以上出現することを採用の目安とした．例えば，4 年間のニュースを解析対象にする場合，約 1000 営業日において出現回数が 100 回以下のキーワードは採用しない．1 年間のニュースを解析対象にする場合は，約 300 営業日において出現回数が 30 回以下のキーワードは採用しない．

3.3 実験

本研究では 2011 年から 2016 年に QUICK ニュース解析サービスから配信されたニュースヘッドラインをテキスト情報として使用した．キーワードの重要性は，流行や経済状況によって時間変化すると考えられるため，重要キーワードを抽出するニュースの参照期間を変えながら，株価に対するボラティリティを高めるキーワードを抽出し，キーワードリストを作成する．本研究で用いるニュースの参照期間として，過去 4 年間または 1 年間のように変更し，抽出される重要キーワードの違いを調査する．

表 1 は過去 4 年間及び過去 1 年間のニュースを参照した場合における重要キーワード上位 10 ワードを示す．結果として，参照期間 4 年間においては「円」や「東証」,「続伸」など金融市場特有のキーワードが多く見られる．一方，参照期間 1 年間においても金融市場特有のキーワードが多く見られるが,「マザーズ」や「東北」などの固有名詞やイベントワードが重要視される傾向にある．つまりニュース参照期間を長期にすることで，金融市場特有の一般的なキーワードが，短期にするとトレンドワードが抽出されやすい．今回は過去 4 年間と過去 1 年間のニュースヘッドラインを参照したが，より期間を短期に (3 カ月や 1 カ月) することで，より顕著にトレンドワードを抽出できると考えられる．また表 1 および表 2 において,「続伸」や「減」のように客観的にもボラティリティに影響すると考えられるキーワードが自動抽出されている．

表 1: 株価ボラティリティに影響を与える上位 10 ワード．参照期間 4 年間 (左列)，参照期間 1 年間 (右列)

2011~2014 年 (選出数 459)		2014 年 (選出数 282)	
キーワード k	t_k	キーワード k	t_k
中国	63.74	東証	50.96
続伸	60.95	輸出	49.73
反発	55.79	値がさ株	47.02
純	55.78	マザーズ	41.51
輸出	55.48	派遣	39.23
減	54.66	円	39.24
大証	54.59	先行	36.70
安心	52.35	格上げ	36.54
円	50.70	米	35.49
超	49.66	社長	34.92

2012~2015 年 (選出数 500)		2015 年 (選出数 268)	
キーワード k	t_k	キーワード k	t_k
円	96.29	東証	52.13
米	71.40	原油	46.74
中国	69.81	三菱	46.28
減	61.01	脱	43.80
大幅	59.82	発売	42.96
大証	59.33	中国	38.58
売る	58.78	純	37.44
原油	56.77	新潟	37.35
純	56.73	インフラ	37.34
反発	55.82	軟調	36.03

2013~2016 年 (選出数 500)		2016 年 (選出数 306)	
キーワード k	t_k	キーワード k	t_k
東証	112.99	東証	70.83
円	101.71	円	55.67
純	64.11	合併	53.41
減	62.42	純	52.46
米	61.78	公表	50.21
原油	61.38	買い戻す	48.30
安	59.27	海運	45.23
中国	57.73	一段	44.61
大幅	56.05	響く	42.05
反発	54.95	安	40.42

4 アクティブリターンの機械学習

4.1 学習方法

前章で作成したキーワードリストや，新規にリアルタイムに得られるニュースヘッドラインや株式価格情報を用いて，f 日後のアクティブリターン (銘柄 c と TOPIX との差) を予測する．現在を d とすると，f 日後のアクティブリターンは次式となる．

$$r_c^*(d+f) = \frac{x_c(d+f) - x_c(d)}{x_c(d)} - \frac{x_t(d+f) - x_t(d)}{x_t(d)} \tag{6}$$

ここで，x_c は銘柄 c の株価，x_t は TOPIX とする．

図 1 に各期間の役割を示す．緑色は 3 章にて抽出したキーワードリストを作成する期間であり，水色はフォワードテストを行うシミュレーション期間である．黒い太枠はキーワードリストを 1 年ごとに更新し，同期間のシミュレーションを行う．本研究では①～④の 4 パターンの期間においてシミュレーションを行う．

機械学習の手順は次の通りである．まず，ニュース

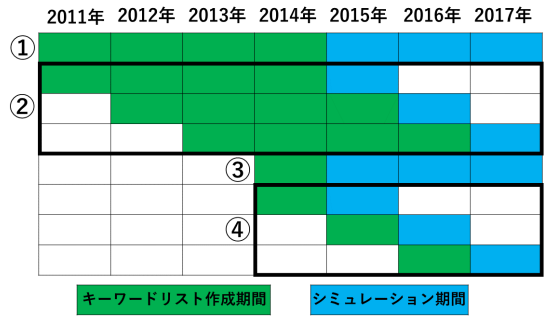


図 1: 各期間の役割 (全 4 パターン)

出現直後のアクティブリターン r_c^* に応じて，各ニュースに教師ラベルを付与する．なお各ニュースに最も関連する銘柄を c とする．

- if $r_c^*(d+f) > 0$, then ニュースの教師ラベル $= 1$
- if $r_c^*(d+f) < 0$, then ニュースの教師ラベル $= -1$
- if $r_c^*(d+f) = 0$, then ニュースの教師ラベル $= 0$

なお $r_c^*(d+f) = 0$ のケースは極めて少ないため，$r_c^*(d+f)$ に関する正・負の 2 クラス判別問題とみなす．

次に，3 章で抽出された重要キーワードを構成要素とする Bag-of-words によって，ニュースヘッドラインをベクトル $\mathbf{v}_c(d)$ として表現する．例えば，d 日にて銘柄 c に関連するニュースヘッドラインが「新作」「株価格付」「黒字」といった重要キーワードを含む時，これらのキーワードに対応するベクトル要素を 1 とし，他を 0 とする．このベクトル $\mathbf{v}_c(d)$ と，将来の $r_c^*(d+f)$ に基づく教師ラベルとの関係性を Support Vector Machine (SVM)[6] によって機械学習する．このように重要キーワードに限定することで，Bag-of-words の高次元化を抑制している．

図 1 のシミュレーション期間において，次のように学習と予測を繰り返す．まず，過去 s 日間 $(d-f-s \sim d-f)$ に配信されたニュースヘッドラインとその教師ラベルを用いて SVM を学習する．次に，当日 d に配信されたニュースヘッドラインを学習済み SVM に入力することで，f 日後のアクティブリターンを予測する．日付 d を進めながら毎日予測を行うが，SVM の再学習は s 日毎とする．なお本研究では，$s = 60$ および $f = 1$ とする．

シミュレーション期間における予測精度を表 2 と表 3 に示す．2 クラス判別問題であるため 50%が正答率の基準となるが，いずれの期間でも基準を超える正答率を示している．しかし $r_c^*(d+f) < 0$ と予測した場合の再現率は高いものの，$r_c^*(d+f) > 0$ と予測した場合の再現率は非常に低い．おそらく相対的にネガティブに作用するニュースが多いため，予測結果が負に偏ったと考えられる．しかし表 3 のように，重要キーワード

表 2: 重要キーワードを 1 つ以上含むニュースヘッドラインに限定した場合

		$r_c^*(d+f) > 0$ と予測			$r_c^*(d+f) < 0$ と予測		
期間	正答率	再現率	適合率	F 値	再現率	適合率	F 値
①	52.8%	1.0%	53.3%	2.0%	99.2%	52.8%	68.9%
②	52.7%	4.9%	48.9%	8.9%	95.4%	52.8%	68.0%
③	52.8%	1.8%	49.6%	3.5%	98.3%	52.9%	68.8%
④	52.6%	7.0%	48.9%	12.2%	93.4%	52.9%	67.5%

表 3: 重要キーワードを 3 つ以上含むニュースヘッドラインに限定した場合

		$r_c^*(d+f) > 0$ と予測			$r_c^*(d+f) < 0$ と予測		
期間	正答率	再現率	適合率	F 値	再現率	適合率	F 値
①	53.1%	5.3%	49.5%	9.6%	95.3%	53.4%	68.4%
②	52.5%	13.4%	48.7%	21.1%	87.4%	53.1%	66.0%
③	52.9%	5.5%	48.4%	9.9%	94.8%	53.2%	68.2%
④	52.5%	14.1%	49.6%	22.0%	87.3%	53.3%	66.2%

数を増やすほど再現率の偏りは改善されている．つまり SVM の学習において，重要キーワードの組み合わせによってニュース全体のボジネガを判断することが重要であり，各キーワードには非線形的な関係があると言える．

5 重要キーワードの妥当性

前章の機械学習においてキーワードの組み合わせが重要であるという知見を得たが，そもそもボラティリティ変化に基づいて重要キーワードを選出する妥当性は不明である．そこで 3 章で抽出された重要キーワードリストを「オリジナルリスト」とし，オリジナルリストに選抜されなかったキーワードをランダムに抽出することで「ランダムリスト」を 100 セット作成する．なおランダムリストに含まれるキーワード数は全てオリジナルリストと同数にする．次にオリジナルリストとランダムリストのそれぞれで Bag-of-words を構成し，各シミュレーション期間において SVM の学習および予測を行った．

それぞれの正答率を図 2〜図 5 に示す．なお，重要キーワードを 3 ワード以上含むニュースヘッドラインに限定した場合である．いずれの期間においても，オリジナルリストによる正答率はランダムリストによる正答率の中央値を超えており，重要キーワードの抽出においてボラティリティ変化に着眼する妥当性を確認

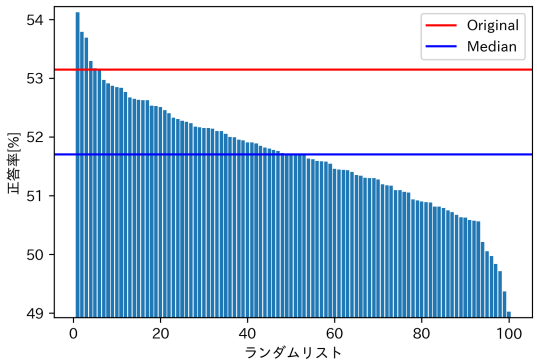


図 2: オリジナルリストによる正答率 (赤線) と ランダムリストによる正答率の中央値 (青線) の比較．重要キーワードを 3 つ以上含むニュースヘッドラインに限定した場合．ただし参照期間①の場合

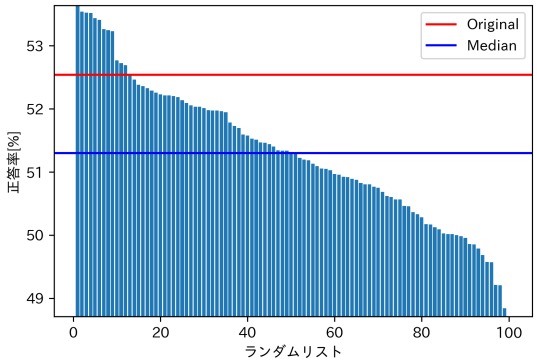


図 3: 図 2 と同様．ただし参照期間②の場合

できる．

6　アクティブ運用への活用

4 章の機械学習において 50%以上の正答率でアクティブリターン $r_c^*(d+f)$ を予測できる結果を得た．そこでアクティブ運用へ活用した場合，どの程度の収益を得られるかシミュレーションを行う．運用方法として，$r_c^*(d+f) > 0$ と予測された銘柄 c を当日 d の終値でロングし，$d+f$ 日の終値で手仕舞いする．$r_c^*(d+f) < 0$ と予測された場合は，銘柄 c を当日 d の終値でショートし，$d+f$ 日の終値で手仕舞いする．なお提案手法の有用性を確認すべく，ランダムにロングまたはショートを行う「ランダム法」とパフォーマンスを比較する．

以下の指標により，運用パフォーマンスを指標する．

累和アクティブリターン G_N

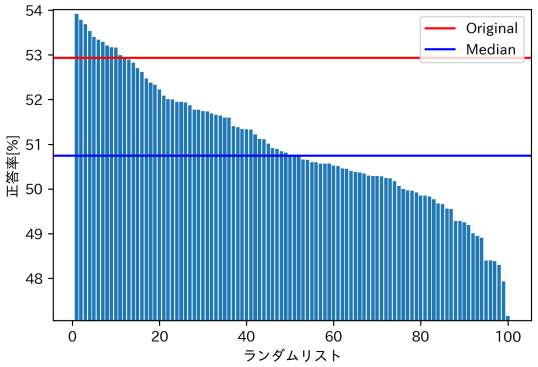


図 4: 図 2 と同様．ただし参照期間③の場合

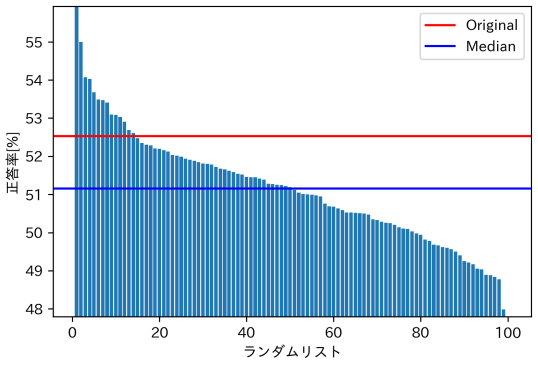


図 5: 図 2 と同様．ただし参照期間④の場合

$$G_N = \sum_{n=1}^{N} a_n r_{c_n}^*(d+f) \times 100\ [\%] \tag{7}$$

ここで，n は重要キーワードを含むニュースヘッドラインのインデックス番号であり，c_n はこのニュースに最も関連する銘柄を表す．運用行動として，$r_{c_n}^*(d+f) > 0$ と予測した場合は $a_n = 1$ (ロング) とし，$r_{c_n}^*(d+f) < 0$ と予測した場合は $a_n = -1$ (ショート) とする．なお，ニュース n に基づいて獲得したアクティブリターンを

$$G_n = a_n r_{c_n}^*(d+f) \tag{8}$$

と書く．

平均アクティブリターン $\bar{G}$

$$\bar{G} = \frac{1}{N}\sum_{n=1}^{N} G_n = \frac{G_N}{N}\ [\%] \tag{9}$$

勝率 W

$$W = \frac{|\{n|G_n > 0\}|}{N} \times 100\ [\%] \tag{10}$$

全運用回数 N に対して，プラス収益を得た運用回数の割合を示す．

インフォメーションレシオ I_R

$$I_R = \frac{\bar{G}}{\sigma_G} \tag{11}$$

ここで σ_G は G_n の標準偏差である．つまり，$\sigma_G = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(G_n - \bar{G})^2}$ である．この I_R が高いほど，リスクに対するリターンが大きい．

プロフィットファクター P_F

$$P_F = \frac{\sum\{G_n | G_n > 0\}}{\sum\{G_n | G_n < 0\}} \tag{12}$$

プラス収益の総和とマイナス収益の総和の比率であり，$P_F > 1$ であれば効率的な投資運用であるといえる．

ペイオフレシオ P_R

$$P_R = \frac{\langle\{G_n | G_n > 0\}\rangle}{\langle\{G_n | G_n < 0\}\rangle} \tag{13}$$

プラス収益の平均値とマイナス収益の平均値の比率であり，$P_R > 1$ であれば 1 取引あたりの利益が損失よりも大きいことを示す．

運用パフォーマンスを表 4 と表 5 に示す．本研究ではニュースヘッドラインに関係する全銘柄を運用対象とするため運用回数が多い．そのため累和アクティブリターン G_N は大きいものの，1 運用あたりの平均アクティブリターン $\bar{G}$ は小さい．また前述と同様に，3 つ以上の重要キーワードを含むニュースヘッドラインに限定した方が勝率 W が高く，I_R や P_F などの指標からも運用の効率性が向上していることを確認できる．

図 6〜図 9 に，累和アクティブリターン G_N の推移を示す．なお重要キーワードを 3 つ以上含むヘッドラインを対象にした場合である．ランダムにロングまたはショートを行う「ランダム法」に比べ，提案手法 (運用コスト $C = 0\%$) の優位性を確認できる．したがって本提案手法による運用判断は妥当であると言える．また，表 5 において勝率 W は 53%程度にも関わらず，運用回数が多いため順調にアクティブリターンを積み上げている様子を確認できる．しかし運用回数が多いほど，運用コストも大きくなる．そこで，式 (7) に運用コスト C を差し引くことで G_N を計算する．

$$G_N = \sum_{n=1}^{N} \left[a_n r^*_{c_n}(d+f) - C\right] \times 100 \ [\%] \tag{14}$$

結果として，運用コストが $C < 0.10\%$ (10bp 未満) であれば超過収益を維持できる可能性がある．

表 4: アクティブ運用において，重要キーワードを 1 つ以上含むヘッドラインに限定した場合

期間	G_N	$\bar{G}$	W	I_R	P_F	P_R
①	2836%	0.036%	52.79%	0.013	1.047	0.94
②	2701%	0.037%	52.64%	0.013	1.049	0.94
③	3024%	0.043%	52.81%	0.015	1.056	0.94
④	2180%	0.035%	52.58%	0.012	1.045	0.94

表 5: 表 4 と同様．ただし重要キーワードを 3 つ以上含むヘッドラインに限定した場合

期間	G_N	$\bar{G}$	W	I_R	P_F	P_R
①	1988%	0.094%	53.16%	0.030	1.113	0.98
②	1491%	0.068%	52.50%	0.023	1.085	0.98
③	1921%	0.104%	52.96%	0.036	1.133	1.01
④	1550%	0.089%	52.79%	0.029	1.108	0.99

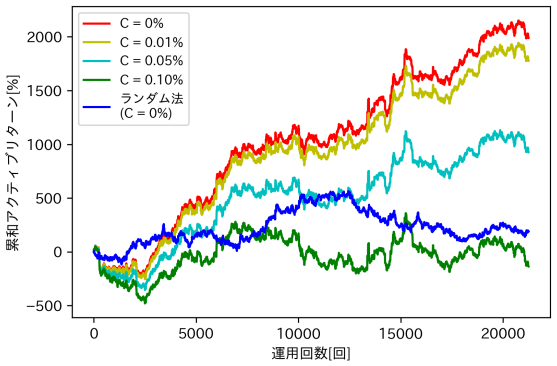


図 6: 参照期間①における累和アクティブリターンの推移．ただし重要キーワードを，3 つ以上含むニュースヘッドラインを対象とした場合

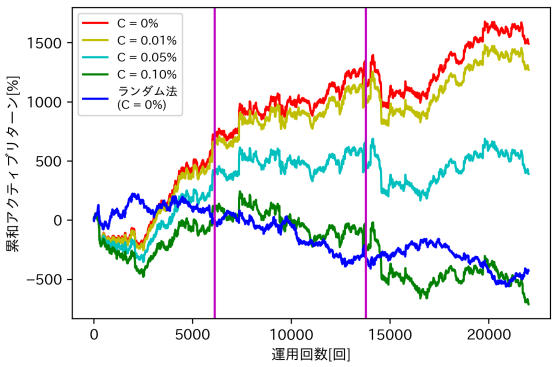


図 7: 図 6 と同様．ただし参照期間②の結果．縦線はキーワードリストの更新タイミング

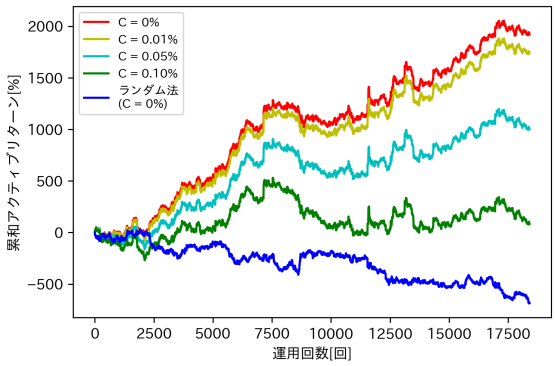


図 8: 図 6 と同様．ただし参照期間③

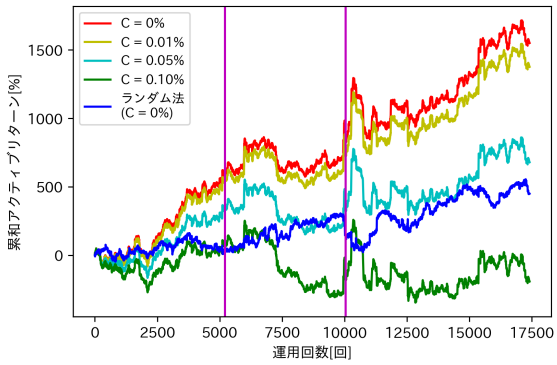


図 9: 図 6 と同様．ただし参照期間④の結果．縦線はキーワードリストの更新タイミング

7　まとめ

本研究ではニュースヘッドラインを解析対象とし，株価のボラティリティを高めるサプライズワードを抽出し，それらの組み合わせと翌日のアクティブリターンの関係を機械学習した．統計的な比較実験を通じてボラティリティーに着眼する妥当性を確認し，アクティブ運用においても 10bp 未満の運用コストならば超過収益を獲得できる可能性を得た．

本研究の一部は，文科省科研費基盤研究（C）(No. 16K00320) の助成により行われました．

参考文献

[1] 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.

[2] 五島圭一, 高橋大志: “株式価格情報を用いた金融極性辞書の作成,” 自然言語処理, vol.24, no.4, pp.547–577, 2017.

[3] C. Kearney and S. Liu: “Textual Sentiment in Finance: A Survey of Methods and Models,” *International Review of Financial Analysis*, vol.33, pp.171–185, 2014.

[4] 株式会社 QUICK: https://corporate.quick.co.jp (参照日 2020.2.23)

[5] MeCab: https://taku910.github.io/mecab/ (参照日 2020.2.23)

[6] C. Cortes and V. Vapni: “Support Vector Networks,” *Machine Learning*, vol.20, no.3 pp.273–297, 1995.

経済因果チェーン検索システムの構築と応用

Economic Causal Chain Search System and its Application

和泉　潔[1]　坂地　泰紀[1]

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 Innovation, School of Engineering, the University of Tokyo

Abstract: In this research, we introduce a system for searching causal relationships of events related to economics and finance in a chain-like manner from a database extracted from economic text data. This system also introduces stock price analysis using this economic causal-chain search.

はじめに

経済ニュース記事を読んでいると，株価の動きや商品の売り上げ，雇用や貿易など様々な経済事象に関する原因と結果の関係の記述がよく出てくる．例えば「将来の少子高齢化が〇〇を引き起こし，それに関連した〇〇に関する需要が増えることが予想される」とか「現在の株価下落は〇〇による市場参加者のリスク警戒心理の高まりを反映している」など，ある事象の経済的な波及効果やその事象の原因について述べている．

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検索する手法を開発する．本手法を用いて，ユーザが入力した語句に対して因果系列を表示し，ユーザが適切な系列に選択したり適切でない系列を削除したりできるシステムを実装した．本システムで作成された因果チェーン，経済的な波及効果と要因列挙が可能となる．

本稿では，因果チェーン検索システムと，本システムやアルゴリズムを用いた応用例を紹介する．

経済的因果関係の抽出とデータベース構築

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[1]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．

因果チェーンの構築アルゴリズム

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には，[2, 3]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. Step 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. Step 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 Step 2, Step 3 を繰り返す.

因果チェーン検索には 2 種類の方向があり，前向きの場合は入力テキストを最初の原因表現として，原因-結果の順で波及効果を検索する(図 1a)．後ろ向きの因果チェーン検索の場合は，入力テキストを最終の結果表現として，結果-原因の順で潜在的な遠因を検索する(図 1b).

図 1. 前向きと後ろ向きの因果チェーン検索

経済因果チェーン検索システム

上述の因果チェーン構築アルゴリズムを基に，ユーザが提示された因果関係を編集できる機能を加えた経済因果チェーン検索システムを実装した [1][4]. 以下に本システムの動作する様子を説明する.

最初に開始テキストをユーザが入力する(図 2). その際に，原因から結果への波及効果の検索か結果から原因への要因検索かを指定する．さらに，因果チェーンを構築する対象となる因果関係が発行された期間を限定することも可能である.

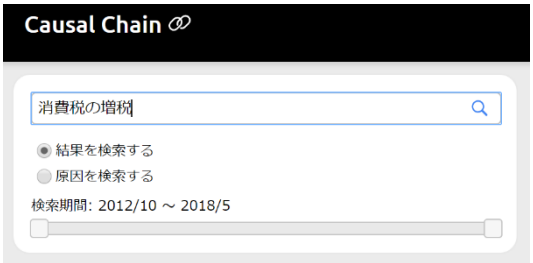


図 2. 開始テキストを入力するテキストボックス

テキストボックスの右にある検索ボタンをクリックすると，開始テキストに連鎖する因果関係を表示する(図 3). デフォルトでは関連性の高い順に 3 個の因果関係を表示する．より多くの因果関係を見たければ，「もっと見る」ボタンをクリックし因果関係ノードの表示を増やすことが出来る．因果関係の各ノードに対して，ユーザが適切な因果でないと判断した場合は，各ノードの右上にある削除ボタンを押して，ノードを削除できる.

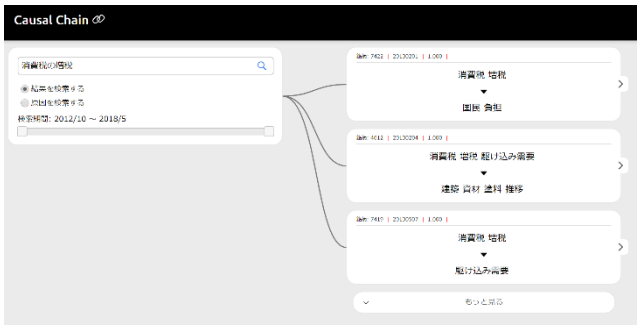


図 3. 連鎖する因果関係の表示

各ノードからさらに因果チェーンを伸ばしたい場合は，各ノードの右側にある「>」ボタンをクリックすると，クリックされたノードを終端ノードとして関連する因果関係が追加される(図 4).

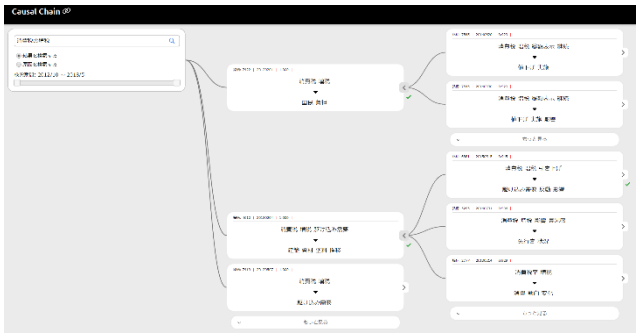


図 4. 因果チェーンの追加

上記の因果チェーンの構築作業を繰り返し，ユーザが必要とする因果チェーンを構築できたら，構築した因果チェーンをファイルに保存することが出来る.

経済因果チェーンを用いた株価分析

経済因果チェーンの検索結果を用いて株価動向分析を行った研究事例を 2 つ紹介する.

因果関係を用いた銘柄間関連性指標

[1] 本システムは 2019 年 3 月より東京大学 和泉研究室のサイト (http://socsim.t.u-tokyo.ac.jp)にて公開予定である.

[6]は，本稿の経済因果チェーンの検索結果を持ちいて企業間の関連性の指標を計算した．ある企業 X の決算短信に出現した因果情報に後ろ向きの因果チェーン検索を行い，X の原因表現に類似する結果表現を有する因果情報を他の企業 Y の決算短信に出現する回数を X と Y の結びつきの強さとした．ある時点で企業 Y の株価に大きな変動があった場合に，市場参加者はその変動要因を類推し，変動要因に関連する企業 X の株式が売買されると仮定した．また Y から X への株価変動の伝播には，要因分析にかかるタイムラグが存在するとした．上記の仮定に基づき，TOPIX500 の構成銘柄に対して，先月の月次リターン(株価変動)が上位 20%の上昇した銘柄と因果関係で関連する銘柄を買った．逆に，先月に下位 20%の下落した銘柄に関連する銘柄を売るという売買を毎月行った．2012 年 12 月から 2019 年 1 月のデータで検証した結果，年率で 6.92%の正の利得が得られた．

因果関係を用いたニュースイベント分析

[7]は，特定のフレーズと関連する企業のリストを，因果チェーンの検索結果を用いて作成した．「小麦価格」という入力テキストに対して前向き因果チェーン検索の結果，図 A のような因果チェーンを得た．各層の因果情報を含む決算短信を発行した企業を調べて，各層で初めて出現した企業のリストを作成した．その結果，先ほどの「小麦 価格」から構築した因果チェーンでは，前向き方向と後ろ向き方向のどちらの検索結果でも，第 1 層では製粉会社や製パン業者など小麦価格に直接関連する企業が出現した．第 2 層、第 3 層には，小麦価格と間接的な関連を持つ企業が出現するようになった．例えば，前向き方向の検索の場合は，化学製品・調味料・食品製造機械メーカーなどが第 2 層と第 3 層で初出となった．後ろ向き方向の場合は，より広範な要因に関連する商社が出現した．このリストに含まれる企業に関して，2018 年 6 月以降で見出しに「小麦 価格」を含む大きなニュース記事が発行された時点の前後 20 営業日の価格変化を調べた．その結果，ニュース発行前に比べ発行後の価格変動が増大していた．さらに，より深い階層で出現した銘柄の株価の方がより大きな変動が長く持続した．

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するシステムの紹介をした．本システムにより，ユーザが入力したイベントからの経済的な波及効果を推論したり，ニュース記事からの因果チェーンをたどった関連銘柄を提示したりするなどのサービス応用のアイデアを紹介した．

謝辞

本システムを実装するにあたって，日名 洋平 氏と松田 健児 氏に多大な助力をいただいた．ここに深謝の意を表する．

参考文献

[1] 坂地 泰紀, 酒井 浩之, 増山 茂: 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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 (to be appeared)

[6] Nakagawa, K.; Sashida, S.; Sakaji, H.; Izumi, K. Economic Causal Chain and Predictable Stock Returns. In Proceedings of the 7th International Conference on Smart Computing and Artificial Intelligence (SCAI 2019), 655-660, 2019.

[7] Kiyoshi Izumi, Shintaro Suda, and Hiroki Sakaji, Economic News Impact Analysis Using Causal-Chain Search from Textual Data, In Proceedings of KDF2020: The AAAI-20 Workshop on Knowledge Discovery from Unstructured Data in Financial Services, 2020.

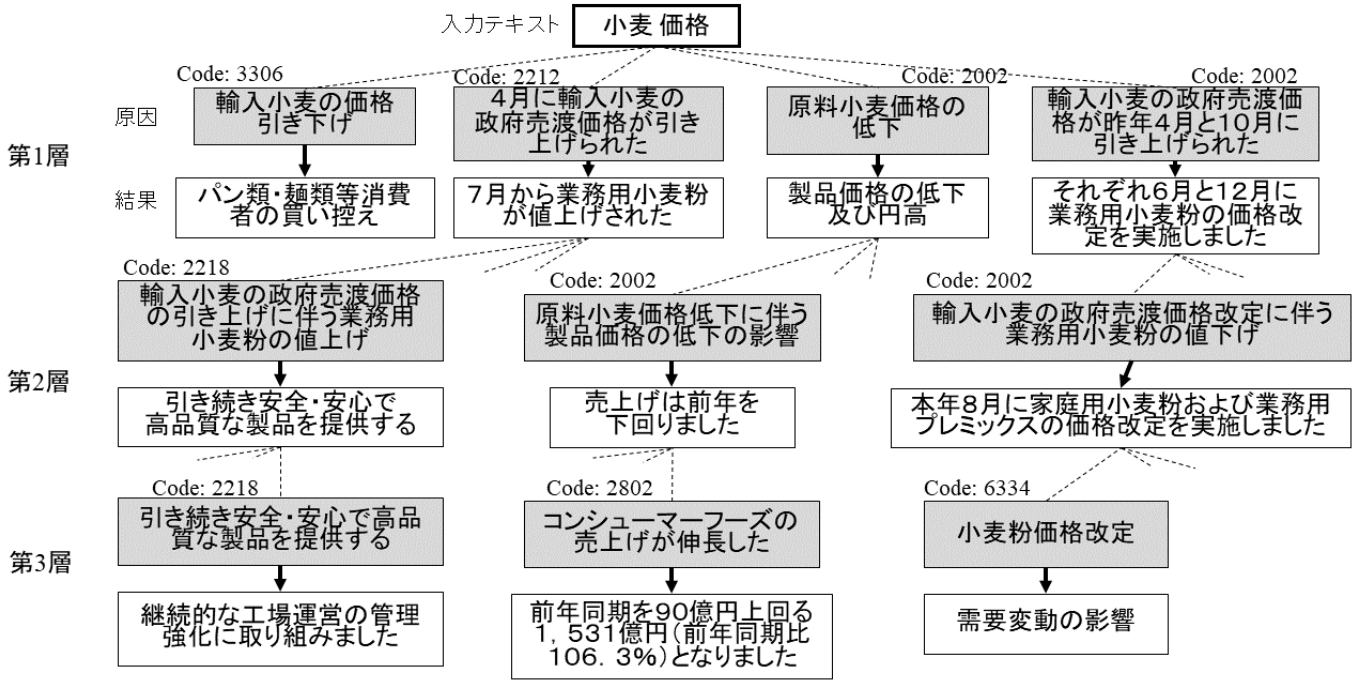


図 A 「小麦 価格」からの前向き因果チェーンの結果の一部

経済因果チェーンを用いたリードラグ効果の実証分析

Empirical Study on Lead-Lag Effect with Economic Casual Chain

中川 慧 [1*]　指田 晋吾 [1] 坂地 泰紀 [2] 和泉 潔 [2]
Kei Nakagawa[1] Shingo Sashida[1] Hiroki Sakaji[2] Kiyoshi Izumi[2]

[1] 野村アセットマネジメント株式会社
[1] Nomura Asset Management Co,Ltd.
[2] 東京大学大学院工学系研究科
[2] The University of Tokyo School of Engineering

Abstract: A lead-lag effect in stock markets describes the situation where one (leading) stock return is cross-correlated with another (lagging) stock return at later times. There are various methods for stock return forecasting based on such a lead-lag effect. One of the most representative methods is based on the supply chain network. In this research, we propose a stock return forecasting method with an economic causal chain. The economic causal chain refers to a cause and effect network structure constructed by extracting a description indicating a causal relationship from the texts of Japanese financial statement summaries. We examine the following lead-lag effect. (1) whether lead-lag effect spreads to the 'effect' stock group when there is a large stock fluctuation in the 'cause' stock group in the causal chain. (2) whether lead-lag effect spreads to the 'cause' stock group when there is a large stock fluctuation in the 'effect' stock group in the causal chain. We confirm the existence of the both side of lead-lag effect and the evidence of stock return predictability across causally linked firms in the Japanese stock market.

1　はじめに

株価の予測可能性は、学術的にも実務的にも重要な研究テーマであり、伝統的なファイナンスの観点の他、様々な分野からの知見も取り入れられ、これまでに株価を予測するために様々な方法が提案されてきた。伝統的なファイナンスにおいては、株式や債券などの金融市場において、資産価格は情報を瞬時に織り込み、超過リターン (投資家が取るリスクに見合うリターンを超すリターン) を得ることはできず、資産価格の予測は不可能とする効率的市場仮説が [10] によって提唱された。ここでは、織り込まれる情報の程度によって、市場の効率性をウィーク型、セミストロング型、ストロング型の 3 種類に分類される。しかし、効率的市場仮説に基づく市場は、現実の市場とは異なる姿であり、これまで効率的市場仮説の下では説明のできないアノマリーが多数観察されている。このようなアノマリーのうち、クロスセクション分析によって明らかになった株価を説明力する属性ないし特徴量を「ファクター」と呼ぶ。ファイナンス分野の長年の実証研究により、ある断面でどのようなファクターを持つ銘柄群が相対的に株価が上昇し、どの銘柄群が下落するかが明らかにされてきた。その代表的なモデルとして、Fama と French[2, 3] による 3 ファクターモデルがある。彼らは株価リターンのクロスセクションの構造は、Beta(市場ポートフォリオに対する感応度)、SMB(時価総額)、HML(PBR の逆数) の 3 つのファクターによって説明できると主張した。Fama-French の 3 ファクターモデルは、それ以前の代表的なモデルであった CAPM[15, 8, 12]) に比べ、モデルの説明力が高いことが、米国を含む様々な株式市場において実証 [4] されており、学術と実務の両方でスタンダードな資産価格モデルの一つとして認識されている。これ以降、Fama-French の 3 ファクターモデル以外のファクターが次々と発見された。その結果、Harvey ら [5] によると、対象を一流学術誌および高い評価を得ているワーキングペーパーに絞っても、2012 年までに累計 300 を超える数多くのファクターが発見された。

このようなクロスセクション分析に基づくファクター以外にも、複数の資産間で観察されるアノマリーの一つにリードラグ効果がある。

リードラグ効果とは、市場に情報が発信された際、ある資産価格に他の資産価格が遅れて反応する現象であ

*連絡先： 野村アセットマネジメント株式会社
〒 103-0027 東京都中央区日本橋 1 丁目 12-1
E-mail: kei.nak.0315@gmail.com

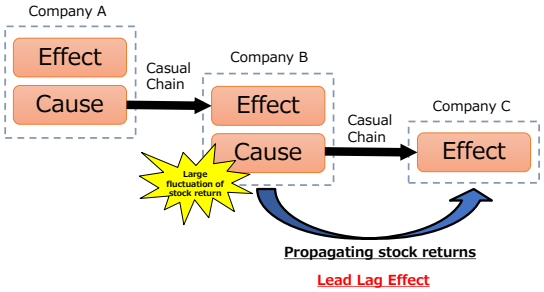


図 1: The illustration of the lead-lag effect with an economic causal chain. In the case of the large fluctuation is propagated to forward.

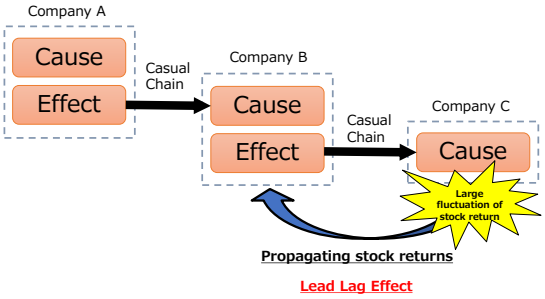


図 2: The illustration of the lead-lag effect with an economic causal chain. In the case of the large fluctuation is propagated to backward.

り、個々の資産によって価格に情報が織り込まれる速度が異なることに起因する現象である。リードラグ効果の代表的な実証分析例として、サプライチェーンを用いてリードラグ効果を計測した、[1] の研究がある。本研究では、サプライチェーンではなく、因果チェーンに基づいたリードラグ効果の計測手法の提案と実証分析を行う。

因果チェーンとは決算短信の文章中から因果関係を示す記述を抽出して作成した因果のネットワーク構造をいう。つまり決算短信から、人間が認知するであろう因果関係を抽出したネットワークである。もし投資家が因果チェーンを適切に認識しており、因果チェーンに基づく情報伝播が即座に反映されるならば、リードラグ効果は存在しないはずである。図 1 は因果チェーンのネットワーク構造を示している。ここで、図中の点線内は各企業の決算短信の文章において、業績にかかわる原因と結果を表す文を表す。Company B と Company C に着目すると、企業 B の結果を表す箇所と企業 C の原因を表す箇所とが紐付くことで因果関係が構築されることがわかる。(以降、文中では Company B に該当するものを結果側、Company C に該当するものを原因側と表記する。) 我々は、因果チェーンを用いたリードラグ効果の検証を、(1) 因果関係を持つ結果側の銘柄群に株価のショックがあったとき、原因側の企業群にその影響が順伝播するかどうか (図 1)、(2) 因果関係を持つ原因側の銘柄群に株価のショックがあったとき、結果側の企業群にその影響が逆伝播するかどうか (図 2) を日本の株式市場を対象に検証する。

2 先行研究

リードラグ効果を検証した研究は多くあるが代表的なものとしてサプライチェーンを用いたリードラグ効果を計測した、[1] の研究がある。米国企業のサプライチェーンから、サプライヤー企業とカスタマー企業の関係性を使った株価予測可能性を調査し、1980 年から 2004 年の検証期間で以下の検証を行った。カスタマー企業にファンダメンタルズや株価の変化が生じたときに、市場が効率的であればその情報は瞬時にサプライヤー企業へ伝播するはずである。カスタマー企業の株価変化が相対的に高かったサプライヤー企業群を買い、カスタマー企業の株価変化が相対的に低かったサプライヤー銘柄群を空売りするカスタマーモメンタム戦略によるポートフォリオを構築した。このロングショートポートフォリオのリスク調整後リターンが有意に正であると報告している。[1] が個別企業のカスタマーサプライヤー情報を財務諸表から収集し、カスタマーモメンタムの計測するのに対し、[11, 14] は業種ベースでサプライヤーカスタマー関係を定義し、カスタマーモメンタムを計測している。同様に、磯貝ら [16] は日本株式市場においてもカスタマーモメンタムは観測されることを報告している。サプライチェーンを前提としない、リードラグ効果を用いた株価予測可能性の研究は米国市場では、 [9] の研究以降多く議論されており、[6] では米株式市場を対象に、同業種内と他業種間でのリードラグ効果を検証した。米株式市場ではリードラグ効果は他業種間においては明確に確認できず、同業種内で強く存在すると方向した。これとは対照的に、他業種間におけるリードラグ効果が [13] によって報告されている。

本研究では、サプライチェーンや市場データではなく、テキスト情報のみから構築された因果チェーンを用いたリードラグ効果の検証を行う (表 1)。

3 因果チェーンの構築

ここでは本研究でリードラグ効果の検証に使用する因果チェーンの構築手法を [7] に基づき概観する。因果チェーンは、因果関係について記述されたテキストデータを解析し，そこから人間が認知するであろう因

表 1: The methods and data sources in lead lag measurement

Author	Method	DataSource
Cohen[1]	企業間サプライチェーン	サプライチェーン (Compustat より取得)
磯貝ら [16]	企業間サプライチェーン	サプライチェーン (有価証券報告書より取得)
Menzly, Shahrur [11, 14]	業種間サプライチェーン	サプライチェーン (CRSP より取得)
Rapach[13]	LASSO 回帰	株価リターン
本研究	因果チェーン	テキストデータ (決算短信より取得)

Input: A list of cause–effect expressions CI
CI_i = (Cause Expression c_i, Effect Expression e_i, Company cp_i, Date d_i)
Output: A list of causal chain LCC

$LCC \leftarrow \emptyset$
for each $(c_i, e_i, cp_i, d_i) \in CI$ **do**
  for each $(c_j, e_j, cp_j, d_j) \in CI$ **do**
    $similarity \leftarrow getSimilarity(e_i, c_j)$
    if $similarity \geq threshold$ **then**
      $LCC \leftarrow LCC + (c_i, e_i, cp_i, d_i, c_j, e_j, cp_j, d_j)$
    end if
  end for
end for
return LCC

図 3: Construction of a causal chain

果関係を抽出する。具体的には以下の 3 ステップにより構築される。

Step 1 SVM を使用して日本の決算短信の要約から因果関係の表現を含む文を抽出する。

Step 2 原因と結果を示す構文パターンを定義し、そこに Step 1 で抽出した文における因果関係を抽出する。

Step 3 抽出された因果関係について、その終端ノード (結果側) の表現と一定の類似性がある他の決算短信の起点ノード (原因側) とを連結し因果チェーンを構築する。(図 3)

また、本研究では因果関係にあるそれぞれの決算短信の発行日と企業名をキーとして、因果の出現数を集計する。集計された数値は因果の強さと解釈することができ、本研究ではその数値変化によるリードラグ効果についても測定を行う。

4 実証分析

4.1 データセットと分析手法

前節の因果チェーンを用いて、リードラグ効果の計測を行う。検証期間は 2012 年 12 月から 2019 年 1 月までである。データセットは TOPIX500 構成銘柄群であり、これは流動性の観点から日本の機関投資家の株式ロングショート戦略のユニバースとして利用されることが多い。

はじめに、(1) 因果関係を持つ結果側の銘柄群に株価のショックがあったとき、原因側の企業群にその影響が順伝播するかどうかを次の手順を毎月末繰り返し検証を行う。

Step 1 因果チェーンの結果側の銘柄群を先月から当月のリターンの大小で 5 分位に分ける。

Step 2 各分位の結果側の銘柄群と紐づく、ある閾値 ($threshold$) 以上の総チェーン数の原因側の銘柄群を特定する。

Step 3 原因側の銘柄群を等ウェイトで保有しパフォーマンスを計測する。ただし、因果チェーンは、投資家が認識できることを保証するため、原因側の決算短信の発行日付はリバランス月から直近 1 年分に該当するもので作成。

次に、(2) 因果関係を持つ原因側の銘柄群に株価のショックがあったとき、結果側の企業群にその影響が逆伝播するかどうかの検証を行う。

Step 1 因果チェーンの原因側の銘柄群を先月から当月のリターンの大小で 5 分位に分ける。

Step 2 各分位の原因側の銘柄群と紐づく、ある閾値 ($threshold$) 以上の総チェーン数の結果側の銘柄群を特定する。

Step 3 結果側の銘柄群を等ウェイトで保有しパフォーマンスを計測する。ただし、因果チェーンは、投資家が認識できることを保証するため、結果側の決算短信の発行日付はリバランス月から直近 1 年分に該当するもので作成。

比較のために、因果チェーンを利用せず単純なリターンの 5 分位ポートフォリオのパフォーマンスも合わせて確認する。

表 2: Annualized return, Risk, and Risk/Return for Quantile portfolios which exploit lead-lag (lead side is effect company) effect using Economic Causal Chains (*Threshold* value from 1 to 10). The test period is from December 2012 to January 2019. In the column labeled Return and row labeled 5-1(Quantile), significant out-performance of one of the two portfolios over the other (in terms of the return) is denoted by asterisks: ∗ denotes significance at the 0.1 level.

Threshold	Quantile	Return[%]	Risk[%]	R/R
1	5-1	0.25	1.14	0.22
2	5-1	0.39	1.06	0.37
3	5-1	-0.19	2.49	-0.07
4	5-1	1.26	4.20	0.30
5	5-1	4.00∗	4.46	0.90
6	5-1	2.51∗	4.21	0.60
7	5-1	5.09∗	8.25	0.62
8	5-1	4.44∗	8.24	0.54
9	5-1	3.65	8.48	0.43
10	5-1	3.33	10.68	0.31

表 3: Annualized return, Risk, and Risk/Return for Quantile portfolios which exploit lead-lag (lead side is cause company) effect using Economic Causal Chains (*Threshold* value from 1 to 10). The test period is from December 2012 to January 2019. In the column labeled Return and row labeled 5-1(Quantile), significant out-performance of one of the two portfolios over the other (in terms of the return) is denoted by asterisks: ∗ denotes significance at the 0.1 level.

Threshold	Quantile	Return[%]	Risk[%]	R/R
1	5-1	0.47∗	0.87	0.54
2	5-1	0.13	1.67	0.08
3	5-1	0.13	2.12	0.06
4	5-1	0.33	3.05	0.11
5	5-1	1.57	4.05	0.39
6	5-1	2.46	5.15	0.48
7	5-1	5.64∗	9.92	0.57
8	5-1	5.50∗	9.74	0.56
9	5-1	5.08∗	12.30	0.41
10	5-1	6.92∗	13.03	0.53

表 4: Annualized return, Risk, and Risk/Return for Quintile portfolios without using Economic Causal Chains.

Quantile	Return[%]	Risk[%]	R/R
1	19.52	18.57	1.05
2	19.46	16.6	1.17
3	18.09	15.58	1.16
4	16.75	15.37	1.09
5	14.99	15.63	0.96
5-1	-4.53	11.52	-0.39

4.2 分析結果と考察

はじめに、(1) 順伝播のリードラグ効果 (結果側から原因側への伝播) の結果は、表 2 の通り。5-1 分位のポートフォリオ (5 分位に含まれる銘柄を等ウェイトで買い、1 分位の銘柄を等ウェイトで売り) のリターンは概ねプラスの傾向となっており、結果側の銘柄のリターンが低調だとその影響が原因側の銘柄群に伝播していることが確認できる。閾値が小さい場合に 5-1 分位のリターンが低い背景としては、原因-結果の紐づけがされる銘柄群が多く、各分位の原因側の銘柄群が重複していることに起因している。閾値 8 以上でリターンが低下している背景としては、原因-結果群の紐付けが数多く存在している銘柄については、その因果関係を認識している投資家も多く存在し、結果側の株価ショック発生時に原因側へ織り込まれる速度が早いことが起因して、検証条件である 1 ヶ月のリバランスではリードラグとしての超過リターンが少なくなったものと考えられる。

次に、(2) 逆伝播のリードラグの結果 (原因側から結果側への伝播) は表 3 の通り。逆伝播の場合、閾値 (Threshold) によらずポートフォリオのリターンはプラスになっている。また閾値 (Threshold) が高くになるにつれて、その傾向は顕著になる。この結果は順伝播 (表 2) の場合と比較すると、逆伝播の場合、株価ショックにおけるラグ側への価格織り込み速度が順伝播よりも遅く、結果的に閾値が高いケースで良好な結果を得られたと考えられる。

また、因果チェーンを使用しない結果は table4 の通りである。この場合、分位が低いとリターンが高く分位が高いとリターンが低くなる、いわゆる短期のリバーサル効果が観測された。チェーン無しでの投資対象となる原因側のユニバースが結果側のユニバースと同等であることを考慮すると、因果チェーンを利用することでリバーサル効果が軽減されているとみることができる。

決算短信は一般に公開されている情報であり、もしも投資家が因果チェーンを適切に認識しており、情報伝播が即座に反映されるならば、リードラグ効果は存在

しないはずである。しかし分析の結果、構築されたポートフォリオにて良好な結果が出たことから因果チェーンを用いたリードラグ効果は存在することが確認できた。つまり、投資家は決算短信から構築された因果チェーンを認識しており、その因果を効率的市場仮説が示唆するように即時に反映せずに、サプライチェーンと同様に時間差をもって反映すること (リードラグ効果) が確認できた。これはモメンタム効果と同様に投資家の過少反応に起因すると考えられる。これは、投資家はニュース等を即座に価格に反映させるのではなく、時間をかけて徐々に織り込んでいくという現象である。情報に対して過少反応するため、理論価格からの乖離が生じて、その乖離が修正される過程で分析結果のようなリードラグ効果が計測されたと考えられる。

なぜ、このような過少反応が存在するかという背景についてはより詳細な検証が必要であるが、次のような投資家の注意力、判断能力の限界に基づく仮説が背後にあると考える。サプライチェーンを用いた先行研究 [1, 16] では、カスタマーモメンタムは観測されるものの、サプライヤーからカスタマー方向への順方向へのモメンタム (サプライヤーモメンタム) は観測されていない。これは、投資家にとって、サプライヤーからカスタマーへの順方向の波及効果を認知する方が、カスタマーからサプライヤーへの逆方向の波及効果を認知するよりも容易であるという背景がある。一方で、本研究では、投資家にとって有価証券報告書等に直接的に記載されるサプライチェーンよりも、推論が必要で認知が困難であると考えられる因果関係に基づく因果チェーンを用いた。そのため、ある企業の株価の変動要因が、当該企業と因果関係にある企業に伝播あるいは逆伝播するという認知が遅れ、順方向、逆方向ともにリードラグ効果が存在したと考えられる。

5 まとめ

本研究では、因果チェーンによって構築されたネットワーク構造を利用したリードラグ効果の計測手法の提案と実証分析を行った。因果チェーンとは決算短信の文章中から因果関係を示す記述を抽出して作成したネットワーク構造をいう。決算短信は一般に公開されている情報であり、もし投資家が因果チェーンを適切に認識おり、情報伝播が即座に反映されるならば、リードラグ効果は存在しないはずである。日本株式市場を対象に、我々は因果チェーンを用いて、(1) 因果関係を持つ結果側の銘柄群に株価のショックがあったとき、原因側の企業群にその影響が順伝播するかどうか、(2) 因果関係を持つ原因側の銘柄群に株価のショックがあったとき、結果側の企業群にその影響が逆伝播するかどうかを検証した。その結果、因果チェーンを用いたリードラグ効果は両方向において存在することが確認できた。

これはモメンタム効果と同様に投資家の過少反応や注意力の限界に起因すると考えられる。今回の分析はユニバースを TOPIX500 に限定し、さらに因果チェーンも限定したが、より広範なユニバースや因果チェーンでの分析が今後の課題である。また、投資家の注意力の限界という観点から、単純な原因と結果の 1 次の繋がりではなく、より高次の因果関係の伝播をより詳細にみていく必要もある。

参考文献

[1] Lauren Cohen and Andrea Frazzini. Economic links and predictable returns. *The Journal of Finance*, Vol. 63, No. 4, pp. 1977–2011, 2008.

[2] Eugene F Fama and Kenneth R French. The cross-section of expected stock returns. *the Journal of Finance*, Vol. 47, No. 2, pp. 427–465, 1992.

[3] Eugene F Fama and Kenneth R French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, Vol. 33, No. 1, pp. 3–56, 1993.

[4] JM Griffin. Are the fama and french factors global or country specific? *Review of Financial Studie*, Vol. 15, No. 3, pp. 783–803, 2002.

[5] Campbell R Harvey, Yan Liu, and Heqing Zhu. … and the cross-section of expected returns. *The Review of Financial Studies*, Vol. 29, No. 1, pp. 5–68, 2016.

[6] Kewei Hou. Industry information diffusion and the lead-lag effect in stock returns. *The Review of Financial Studies*, Vol. 20, No. 4, pp. 1113–1138, 2007.

[7] Kiyoshi Izumi and Hiroki Sakaji. Economic causal-chain search using text mining technology. In *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pp. 61–65, Macao, China, 8 2019.

[8] John Lintner. The valuation of risk assets on the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, Vol. 47, pp. 13–37, 1965.

[9] Andrew W Lo and A Craig MacKinlay. When are contrarian profits due to stock market overreaction? *The review of financial studies*, Vol. 3, No. 2, pp. 175–205, 1990.

[10] Burton G Malkiel and Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, Vol. 25, No. 2, pp. 383–417, 1970.

[11] Lior Menzly and Oguzhan Ozbas. Cross-industry momentum. 2006.

[12] Jan Mossin. Equilibrium in a capital asset market. *Econometrica*, Vol. 34, No. 4, pp. 768–783, 1966.

[13] David E Rapach, Jack Strauss, Jun Tu, and Guofu Zhou. Industry interdependencies and cross-industry return predictability. 2015.

[14] Husayn Shahrur, Ying L Becker, and Didier Rosenfeld. Return predictability along the supply chain: the international evidence. *Financial Analysts Journal*, pp. 60–77, 2010.

[15] William F Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, Vol. 19, No. 3, pp. 425–442, 1964.

[16] 磯貝明文, 川口宗紀, 小林寛司. サプライヤー・カスタマーのつながりに基づくクロスモメンタムの株価予測可能性. 現代ファイナンス, Vol. 40, pp. 25–48, 2019.

トピック埋め込み回帰モデルを用いた株価予測

Predicting Stock Prices using Topic Embedding Regression Model

許 蔚然 [1*]　江口 浩二 [2†]
Weiran Xu[1]　Koji Eguchi[2]

[1] 神戸大学 大学院 システム情報学研究科
[1] Graduate School of System Informatics, Kobe University
[2] 広島大学 情報科学部
[2] School of Informatics and Data Science, Hiroshima University

Abstract: In this paper, we aim to predict stock prices by analyzing text data in financial articles. TopicVec is a topic embedding model that represents latent topics in a word embedding space. Here, word embedding maps words into a low-dimensional continuous embedding space by exploiting the local word collocation patterns in a small context window. On the other hand, topic modeling maps documents onto a low-dimensional topic space. Using the topic embedding model, topics underlying each document can be mapped into the word embedding space by combining word embedding and topic modeling.

The topic embedding model has not been used to address regression problem and also has not been used to predict stock prices by analyzing financial articles, to our knowledge. In this paper, by extending the topic embedding model to regression, we propose a topic embedding regression model called TopicVec-Reg to jointly model each document and a continuous label associated with the document. Our method takes financial articles as documents, each of which is associated with a stock price return as a continuous label, so that we can predict stock price returns for new unlabeled financial articles.

We evaluate the effectiveness of TopicVec-Reg through experiments in the task of stock return prediction using news articles provided by Thomson Reuters and stock prices by the Tokyo Stock Exchange. The result of closed test shows that our method brought meaningful improvement on prediction performance.

1　はじめに

金融市場で，人々が経済指標や企業に関するニュースや世の中の出来事などを考慮したうえで投資先を決定するのが一般的である．しかし，これらの情報は常に更新されていて，人の手ですべての情報を把握することは不可能であると考えられる．そこで，人々の負担軽減と意思決定を補佐するため，情報科学の手法でこれらの情報から株価などの金融指標を予測する研究が行われている．本稿では，金融記事のテキストデータを分析し，株価を予測する課題に取り組む．

テキストデータを分析するために，トピックモデルを使うのが一般的である．トピックモデルは文書コレクションに隠れた意味的な構造を発見する統計ツールであり，マーケティング，社会学，政治科学などの分野で応用されている．そのうち多くのトピックモデルは LDA(Latent Dirichlet Allocation) [1] という階層的な確率的モデルをベースとして拡張されてきた．LDA は，文書全体にわたる単語の共起パターンに従って，単語をトピックにグループ化する．コーパスが十分に大きい場合，共起パターンは単語間の意味的関連性をは反映するため，一貫性のあるトピックを発見できる．単語の出現確率はその背後にあると仮定されている潜在トピックによって決まる．そこで，単語は「one hot」（1 つの次元のみ 1 で他は 0 であるような語彙数次元ベクトル）で表現される．しかし，この単純な表現は高次元かつ非常にスパースであるため，与えられたコレクションが巨大な語彙量を持つと，LDA の性能が落ちてしまう [2].

*連絡先： 神戸大学 大学院 システム情報学研究科
〒 657-8501 兵庫県神戸市灘区六甲台町 1-1
E-mail: weland@cs25.scitec.kobe-u.ac.jp
†連絡先： 広島大学 情報科学部
〒 739-8527 東広島市鏡山 1-4-1
E-mail: eguchi@acm.org

単語の「one hot」表現を避けるために，単語埋め込み（word embedding）が提案された [3]．単語埋め込みは，単語の局所的な共起パターンを利用することにより，意味的又は構文的な関連性を符号化された単語の分散表現を学習できる．そのような分散表現を持つ単語のうち，似たような意味を持っていれば，埋め込み空間上で近い場所に位置する．

単語の集合である文書に着目し，各単語の埋め込みが与えられたとすると，文書を「bag-of-vectors」と見なすことができる．関連性のある単語は似た方向に向き，意味的なクラスタができてその重心を意味的重心と見なせる．特定のドメインの文書集合において，いくつかの意味的なクラスタが多くの文書の中に現れる場合があるため，それらを LDA のトピックと見なすことができる．ただし，意味的クラスタの重心は埋め込み空間に存在する．この考え方に基づいて，生成的単語埋め込みモデル PSDVec [7] にトピックを加えることで，トピック埋め込みモデル TopicVec [6] が提案された．TopicVec では，各単語の生成確率は局所的な文脈と大域的なトピックに影響され，トピック埋め込みベクトルは意味的重心として単語埋め込みと同じ埋め込み空間上に存在する．LDA と同様に Dirichlet 事前分布を用いてトピック分布を正規化する．単語の生成においては LDA で使われるカテゴリカル分布の代わりに，PSDVec でのリンク関数にトピックを加えたものを使うと仮定する．これによって，単語間の意味的関連性は埋め込み空間内に余弦類似度として符号化される．TopicVec は多クラステキスト分類タスクにおいて，LDA より優れた性能をもたらすことが報告されている [6]．

TopicVec のようなトピック埋め込みモデルを回帰問題に用いた研究や，それにより金融記事記事のテキストデータから株価を予測する研究は我々の知る限り未だ行われていない．そこで，本稿は，テキストデータとそれに対応する連続値ラベルとの関係をモデル化する回帰機能を付与したトピック埋め込みモデルとして TopicVec-Reg を提案する．TopicVec-Reg での回帰機能は，Supervised Latent Dirichlet Allocation（sLDA）[4] に倣い，1 文書に対応する連続値ラベルをガウス分布に従うと仮定する．さらに，ある企業に関する金融記事をテキストデータ，その企業に対応する株価リターン率を連続値ラベルと見なすことで，金融記事から株価リターン率を予測する．学習過程は変分推論により，回帰係数と潜在トピックを同時に学習する．

本稿の提案手法 TopicVec-Reg の予測性能を評価するために，回帰係数を考慮しない TopicVec を比較対象として，各記事が発信された前後の株価リターン率を予測するクローズドテストを行う．

2 関連研究

Das ら [5] は LDA を単語埋め込みを利用するように拡張し，GaussianLDA を提案した．この手法は，離散的なテキストを連続的な埋め込みの観測に変換し，LDA を実数値データを生成するように適合させる．トピック内の単語はトピック埋め込みを期待値パラメータとする多変量ガウス分布からランダムにサンプリングされると仮定される．しかし，埋め込み間の意味的関連性を測るにはユークリッド距離が最適な選択肢とは限らないことから，これはやや強い仮定であると考えられる．

Dieng ら [2] は LDA と単語埋め込みを結合した生成的モデル ETM を提案した．単語はカテゴリカル分布から生成されると仮定される．そのカテゴリカル分布のパラメータは単語埋め込みと与えられたトピック埋め込みの内積である．ETM は巨大な文書コーパスにおいても解釈可能性のあるトピックを学習できる．さらに，文書データのストップワードを除外しなくても，良いトピック品質を保つことができる．前述の TopicVec は ETM と同じ発想を持っているが，単語はリンク関数によって生成されると仮定される．

Li ら [7] は TopicVec の前身である生成的単語埋め込み方法 PSDVec を提案した．PSDVec は，注目単語と近くにある文脈単語との共起パターンを用いて注目単語を低次元連続的埋め込み空間に写像する．文脈が与えられた単語の条件付き分布は，独立した対数双線形項に因子分解できると仮定される．単語埋め込みと回帰残差をガウス事前分布を仮定して正規化することによって過適合を防ぐ．TopicVec が PSDVec と異なる点は，TopicVec での単語の条件付き分布は，文脈単語だけではなく，トピック埋め込みによる影響も受けることにある．

Blei ら [4] はラベル付けの文書を対象とする教師ありトピックモデル sLDA を提案した．sLDA を LDA に各文書に紐づける応答変数を付け加えたものであると考えられる．応答変数は一般化線形モデルで生成されると仮定され，文書に割り当てられたトピックの期待値をその線形回帰の説明変数とする．文書と応答変数を同時にモデル化することで，新たに与えられた文書の応答変数を良く予測できる潜在トピックを推定できると期待される．この考え方に従い，本稿はトピック埋め込みモデル TopicVec に線形回帰モデルを付与することで，トピック埋め込み回帰モデル TopicVec-reg を提案する．

3 トピック埋め込み回帰モデル

本稿において，語彙サイズが V である D 文書のコーパスを考える．文書 d_i の j 番目の単語を $w_{ij} \in \{1, \cdots, V\}$ で表す．

3.1 TopicVec

本節では文献 [6] に示された TopicVec の概要を示す．まず，前提となる単語埋め込みモデル PSDVec [7] のリンク関数は着目する単語の埋め込み $\boldsymbol{v}_{w_c}$ と文脈単語の埋め込み $\boldsymbol{v}_{w_l}$ を接続する：

$$\begin{aligned} &P\left(w_c | w_0 : w_{c-1}\right) \\ \approx &P\left(w_c\right) \exp\left\{ \boldsymbol{v}_{w_c}^\top \sum_{l=0}^{c-1} \boldsymbol{v}_{w_l} + \sum_{l=0}^{c-1} a_{w_l w_c} \right\}. \end{aligned} \tag{1}$$

ここで，$a_{w_l w_c}$ はバイグラム残差（バイアス）と呼ばれ，$\boldsymbol{v}_{w_c}^\top \boldsymbol{v}_{w_l}$ が捕らえられない非線形部分を示す．

本稿の提案手法のベースである TopicVec の単語の条件付き分布は，PSDVec に基づいて，トピックを文脈単語と見なす：

$$\begin{aligned} &P\left(w_c | w_0 : w_{c-1}, z_c, d_i\right) \\ \approx &P\left(w_c\right) \exp\left\{ \boldsymbol{v}_{w_c}^\top \left(\sum_{l=0}^{c-1} \boldsymbol{v}_{w_l} + \boldsymbol{t}_{z_c} \right) \right. \\ &\left. + \sum_{l=0}^{c-1} a_{w_l w_c} + r_{z_c} \right\}. \end{aligned} \tag{2}$$

ここで，$\boldsymbol{t}_{z_c}$ は注目単語に割り当てられたトピックの埋め込みであり，r_{z_c} はトピックに関する残差である．

TopicVec の生成過程において，単語埋め込み $\boldsymbol{v}_{s_i}$ と残差 $a_{s_i s_j}$ はそれぞれガウス分布から抽出されると仮定される．簡単のために，それらの生成過程を省略し，トピック埋め込みの生成過程は次のようである：

1. トピック k について，半径 γ の超球から一様にトピック埋め込みを抽出する．すなわち $\boldsymbol{t}_k \sim \mathrm{Unif}(B_\gamma)$；
2. 各文書 d_i について:
   (a) ディリクレ事前確率 $\mathrm{Dir}(\boldsymbol{\alpha})$ から混合比率 $\boldsymbol{\phi}_i$ を抽出する；
   (b) j 番目の単語について:
      i. カテゴリカル分布 $\mathrm{Cat}(\boldsymbol{\phi}_i)$ からトピック割り当て z_{ij} を抽出する．
      ii. $P\left(w_{ij} | w_{i,j-c} : w_{i,j-1}, z_{ij}, d_i\right)$ に従って単語 w_{ij} を語彙集合 $\boldsymbol{S}$ から抽出する．

文書データ $\boldsymbol{D}$，単語埋め込み $\boldsymbol{V}$，バイグラム残差 $\boldsymbol{A}$，トピック埋め込み $\boldsymbol{T}$，トピック割り当て $\boldsymbol{Z}$，トピック分布 ϕ に関する対数尤度は以下のようになる：

$$\begin{aligned} &\log p(\boldsymbol{D}, \boldsymbol{A}, \boldsymbol{V}, \boldsymbol{Z}, \boldsymbol{T}, \boldsymbol{\phi} | \boldsymbol{\alpha}, \gamma, \boldsymbol{\mu}) \\ =&C_0 - \log \mathcal{Z}(\boldsymbol{H}, \boldsymbol{\mu}) - \|\boldsymbol{A}\|_{f(\boldsymbol{H})}^2 - \sum_{i=1}^{W} \mu_i \left\|\boldsymbol{v}_{s_i}\right\|^2 \\ &+ \sum_{i=1}^{M} \left\{ \sum_{k=1}^{K} \log \phi_{ik} \left(m_{ik} + \alpha_k - 1\right) \right. \\ &+ \sum_{j=1}^{L_i} \left(r_{i,z_{ij}} + \boldsymbol{v}_{w_{ij}}^\top \left(\sum_{l=j-c}^{j-1} \boldsymbol{v}_{w_{il}} + \boldsymbol{t}_{z_{ij}} \right) \right. \\ &\left. \left. + \sum_{l=j-c}^{j-1} a_{w_{il} w_{ij}} \right) \right\}. \end{aligned} \tag{3}$$

対数尤度を最大化するために，ハイパーパラメータ $\boldsymbol{\alpha}, \gamma, \boldsymbol{\mu}$ が与えられたとし，最適な $\boldsymbol{V}, \boldsymbol{T}, p(\boldsymbol{Z}, \boldsymbol{\phi} | \boldsymbol{D}, \boldsymbol{A}, \boldsymbol{V}, \boldsymbol{T})$ を推定する：

step1 PSDVec で $\boldsymbol{V}^*, \boldsymbol{A}^*$ を最適化する

step2 $\boldsymbol{V}^*, \boldsymbol{A}^*$ を対数尤度関数に代入し，$\boldsymbol{T}^*, p(\boldsymbol{Z}, \boldsymbol{\phi} | \boldsymbol{D}, \boldsymbol{A}, \boldsymbol{V}, \boldsymbol{T})$ を最適化する

step2 で事後分布を解析的に計算するのは不可能であるため，変分分布 $q(\boldsymbol{Z}, \boldsymbol{\phi}; \boldsymbol{\pi}, \boldsymbol{\theta}) = q(\boldsymbol{\phi}; \boldsymbol{\theta}) q(\boldsymbol{Z}; \boldsymbol{\pi})$ で近似する．ここで，KL ダイバージェンスを導入し，推定タスクを変分下限 $\mathcal{L}(q, \boldsymbol{T})$ の最大化に置き換える：

$$\begin{aligned} &\mathrm{KL}(q \| p) \\ &= \log p(\boldsymbol{D} | \boldsymbol{T}) - \left(E_q[\log p(\boldsymbol{D}, \boldsymbol{Z}, \boldsymbol{\phi} | \boldsymbol{T})] + \mathcal{H}(q) \right) \\ &= \log p(\boldsymbol{D} | \boldsymbol{T}) - \mathcal{L}(q, \boldsymbol{T}) \end{aligned} \tag{4}$$

ここで $p = p(\boldsymbol{Z}, \boldsymbol{\phi} | \boldsymbol{D}, \boldsymbol{T})$，$\mathcal{H}(q)$ は q のエントロピーである．変分下限 $\mathcal{L}(q, \boldsymbol{T})$ は

$$\begin{aligned} &\mathcal{L}(q, \boldsymbol{T}) \\ =&\sum_{i=1}^{M} \left\{ \sum_{k=1}^{K} \left(\sum_{j=1}^{K} \pi_{ij}^k + \alpha_k - 1 \right) \left(\psi\left(\theta_{ik}\right) - \psi\left(\theta_{i0}\right) \right) \right. \\ &\left. + \mathrm{Tr}\left(\boldsymbol{T}_i^\top \sum_{j=1}^{L_i} \boldsymbol{v}_{w_{ij}} \boldsymbol{\pi}_{ij}^\top \right) + \boldsymbol{r}_i^\top \sum_{j=1}^{L_i} \boldsymbol{\pi}_{ij} \right\} \\ &+ \mathcal{H}(q) + C_1 \end{aligned} \tag{5}$$

であり，$\mathcal{L}(q, \boldsymbol{T})$ を最大化する最適な q と $\boldsymbol{T}$ を見つけるため，GEM（Generalized Expectation-Maximization）アルゴリズムを用いる：

E-Step：

$$\pi_{ij}^{k} \propto \exp\left\{\psi\left(\theta_{ik}\right) + \boldsymbol{v}_{w_{ij}}^{\top}\boldsymbol{t}_{ik} + r_{ik}\right\}, \tag{6}$$

$$\theta_{ik} = \sum_{j=1}^{L_i} \pi_{ij}^{k} + \alpha_k. \tag{7}$$

M-Step：

$$\boldsymbol{T}_i^{(l)} = \boldsymbol{T}^{(l-1)} + \lambda\left(l, L_i\right) \frac{\partial \mathcal{L}\left(q^{(l)}, \boldsymbol{T}\right)}{\partial \boldsymbol{T}_i}. \tag{8}$$

推定過程のより詳細については文献 [6] に参照されたい．

3.2 TopicVec-Reg

本節では，前節で述べた TopicVec に回帰機能を加えた提案モデル TopicVec-Reg について述べる．TocpicVec-Reg では，各文書に連続値ラベルが紐づけられ，そのラベルはガウス分布に従ってサンプリングされると仮定する．

TopicVec における文書 d_i の生成過程に，以下のような連続値ラベル（応答変数）の生成過程を加える：

3. 連続値ラベル $y|z_{i1:iN}, \boldsymbol{\eta}, \delta^2 \sim \mathrm{N}(\boldsymbol{\eta}^{\top}\bar{\boldsymbol{z}}, \delta^2)$.

ここで用いられるガウス分布の平均は回帰係数 $\boldsymbol{\eta}$ と説明変数であるその文書のトピックの期待値 $\bar{z}$ の内積である．

文書における各単語にトピックが与えられたとき，y の対数尤度の期待値は以下のようになる：

$$\begin{aligned}
&\mathrm{E}\left[\log p\left(y|\boldsymbol{Z}_{1:L_i}, \eta, \delta^2\right)\right] \\
=& -\frac{1}{2}\log\left(2\pi\delta^2\right) \\
&-\frac{1}{2\delta^2}\left(y^2 - 2y\boldsymbol{\eta}^{\top}\mathrm{E}[\bar{\boldsymbol{Z}}] + \boldsymbol{\eta}^{\top}\mathrm{E}\left[\bar{\boldsymbol{Z}}\bar{\boldsymbol{Z}}^{\top}\right]\boldsymbol{\eta}\right)
\end{aligned} \tag{9}$$

ここで，

$$\mathrm{E}[\bar{\boldsymbol{Z}}] = \bar{\boldsymbol{\pi}}_i := \frac{1}{L_i}\sum_{j=1}^{L_i}\boldsymbol{\pi}_{ij}, \tag{10}$$

$$\begin{aligned}
&\mathrm{E}\left[\bar{\boldsymbol{Z}}\bar{\boldsymbol{Z}}^{\top}\right] \\
=&\frac{1}{{L_i}^2}\left(\sum_{j=1}^{L_i}\sum_{m\neq j}\boldsymbol{\pi}_{ij}\boldsymbol{\pi}_{im}^{\top} + \sum_{j=1}^{L_i}\mathrm{diag}\left\{\boldsymbol{\pi}_j\right\}\right)
\end{aligned} \tag{11}$$

である．

TopicVec をベースにしているため，TopicVec-Reg の推定過程において，目的関数 $\mathcal{L}(q, \boldsymbol{T})$ は (9) 式を (5) 式に加えることで得られる：

$$\begin{aligned}
&\mathcal{L}(q, \boldsymbol{T}) \\
=&\sum_{i=1}^{M}\left\{\sum_{k=1}^{K}\left(\sum_{j=1}^{L_i}\boldsymbol{\pi}_{ij}^{k} + \alpha_k - 1\right)\left(\psi\left(\theta_{ik}\right) - \psi\left(\theta_{i0}\right)\right)\right. \\
&+\left(-\frac{1}{2}\log(2\pi\delta^2) - \frac{y_i^2}{2\delta^2}\right) \\
&+\mathrm{Tr}\left(T_i^{\top}\sum_{j=1}^{L_i}\boldsymbol{v}_{w_{ij}}\boldsymbol{\pi}_{ij}^{\top}\right) + \left(\boldsymbol{r}_i^{\top} + \frac{y_i\boldsymbol{\eta}^{\top}}{L_i\delta^2}\right)\sum_{j=1}^{L_i}\boldsymbol{\pi}_{ij} \\
&\left.+\left(-\boldsymbol{\eta}^{\top}\frac{1}{2{L_i}^2\delta^2}\left(\sum_{j=1}^{L_i}\sum_{m\neq j}^{L_i}\boldsymbol{\pi}_{ij}\boldsymbol{\pi}_{im}^{\top} + \sum_{j=1}^{L_i}\mathrm{diag}\{\boldsymbol{\pi}_{ij}\}\right)\boldsymbol{\eta}\right)\right\} \\
&+\mathcal{H}(q) + C_1
\end{aligned} \tag{12}$$

π_{ij}^{k} の更新式を求めるには，$\mathcal{L}(q, \boldsymbol{T})$ から j 番目の単語がトピック k に割り当てられる確率 π_{ij}^{k} を含む項を抽出し，π_{ij}^{k} に関する偏微分を解く．更新式は以下のようになる：

$$\begin{aligned}
\pi_{ij}^{k} \propto \exp\Bigg\{&\psi(\theta_{ik}) + \boldsymbol{v}_{w_{ij}}^{\top}\boldsymbol{t}_{ik} + r_{ik} + \frac{y_i\eta^{k}}{L_i\delta^2} \\
&- \frac{\boldsymbol{\eta}^{\top}\boldsymbol{\Pi}_{i,-j}^{(k)}\boldsymbol{\eta} + (\eta^{k})^2}{2{L_i}^2\delta^2}\Bigg\}.
\end{aligned} \tag{13}$$

ここで，$\boldsymbol{\Pi}_{i,-j}^{(k)}$ は $\sum_{j=1}^{L_i}\sum_{m\neq j}^{L_i}\boldsymbol{\pi}_{ij}\boldsymbol{\pi}_{im}^{\top}$ を π_{ij}^{k} に関して偏微分したものであり，

$$\begin{aligned}
\boldsymbol{\Pi}_{i,-j}^{(k)} :=& \sum_{m\neq j}^{L_i}\boldsymbol{\Pi}_{im}\,\mathrm{diag}\{0^{(1)}, \cdots, 1^{(k)}, \cdots, 0^{(K)}\} \\
&+\mathrm{diag}\{0^{(1)}, \cdots, 1^{(k)}, \cdots, 0^{(K)}\}\sum_{m\neq j}^{L_i}\boldsymbol{\Pi}_{im}
\end{aligned}$$

と定義する．$\boldsymbol{\Pi}_{im}$ は $K \times K$ の行列であり，各行は $(\pi_{im}^{1}, \pi_{im}^{2}, \cdots, \pi_{im}^{K})$ である．

θ_{ik} は（7）式の通りに更新する．

$\mathcal{L}(q, \boldsymbol{T}^{(l-1)})$ のそれぞれ $\boldsymbol{\eta}$ と δ^2 に関する偏微分を解くと，sLDA [4] と同じく $\boldsymbol{\eta}$ と δ^2 の更新式を導出できる：

$$\hat{\boldsymbol{\eta}}_{\mathrm{new}} \leftarrow \left(\mathrm{E}\left[A^{\top}A\right]\right)^{-1}\mathrm{E}\left[A\right]^{\top}\boldsymbol{y} \tag{14}$$

$$\hat{\delta}^2_{\mathrm{new}} \leftarrow \frac{1}{D}\left\{\boldsymbol{y}^{\top}\boldsymbol{y} - \boldsymbol{y}^{\top}\mathrm{E}[A]\left(\mathrm{E}\left[A^{\top}A\right]\right)^{-1}\mathrm{E}[A]^{\top}\boldsymbol{y}\right\} \tag{15}$$

ここで，

$$\mathrm{E}\left[A\right] = \frac{1}{L_i}\sum_{j=1}^{L_i}\boldsymbol{\pi}_{ij}$$

$$\mathrm{E}\left[A^\top A\right] = \sum_{i=1}^{D}\left(\frac{1}{{L_i}^2}\left(\sum_{j=1}^{L_i}\sum_{m\neq j}^{L_i}\boldsymbol{\pi}_{ij}\boldsymbol{\pi}_{im}^\top + \sum_{j=1}^{L_i}\mathrm{diag}\{\boldsymbol{\pi}_{ij}\}\right)\right)$$

である．A は各行が $\bar{\boldsymbol{Z}}_d^\top$ である $D\times(K+1)$ の行列である．$K+1$ 列目は回帰係数のバイアス項に対応する．

$\boldsymbol{T}$ は（8）式の通りに勾配降下法で更新する．

4 評価実験

TopicVec-Reg の予測性能を評価するために，TopicVec を比較対象として，実験を行った．

4.1 データセット

一定とされる単語埋め込み $\boldsymbol{V}^*$ と単語埋め込み残差 $\boldsymbol{A}^*$ は日本語版 Wikipedia を用いた PSDVec で学習済みのものにする．入力されるテキストデータとして，トムソン・ロイター社が 2013 年から 2017 年に発信した日本語金融記事を用いる．各記事に紐づけられた連続値ラベルは株価リターン率を，

$$\frac{\text{記事が出た翌日の終値} - \text{記事が出た前日の終値}}{\text{記事が出た前日の終値}}$$

と設定する．株価は東京証券取引所の歩み値データを利用する．

2ヶ月分の記事を学習用データとする．このようなデータを 3 セットを用意し，それぞれ独立で実験を行った．データセットの概要を表 1 に示す．

表 1: データセット

	data
set1	Nov. 2016~Dec. 2016 # of documents: 931
set2	Jan. 2017~Feb. 2017 # of documents: 912
set3	Mar. 2017~Apr. 2017 # of documents: 845

データに対して前処理を行った．テキストデータに対して，記事の中にある扱いにくい表や必要のない表現の除去した後，MeCab と新語や固有表現を取り揃えた辞書 mecab-ipadic-NEologd [8] [9] [10] を用いて形態素解析を行い，助詞や接続詞などのストップワードを「*」に置き換えた．記事の中に複数の企業が言及された場合，記事を複製してリターン率の大きい企業に残した．最後に，5 文書未満しか出現しない低頻度語と 50 単語未満の文書を除去した．

4.2 実験設定

モデル

TopicVec-Reg: 回帰パラメータの学習をトピック推定と同時に行う

TopicVec+LR: TopicVec でトピック推定した後に線形回帰を行う

TopicVec+LR を比較対象として TopicVec-Reg の評価実験を行った．TopicVec+LR では，トピック推定を TopicVec で行った後，ライブラリ scikit-learn の線形回帰モデル Linear Regression で回帰係数を推定する．

まず学習用データでトピック数が $K=\{5,10,15,20,25\}$ の 5 通りのモデル構築を行い，トピック埋め込み $\boldsymbol{T}$ と回帰係数 $\boldsymbol{\eta}$ を得る．

学習開始時の $\boldsymbol{T}$ の初期値はガウス分布に従う乱数で与え，変分パラメータ $\boldsymbol{\pi}$ に対しては一様分布に従う乱数を与えた．ハイパーパラメータ $\boldsymbol{\alpha}=(0.1,\cdots,0.1)$ とする．

テストを行うときに以下の式に従い連続値ラベルを予測する：

$$\boldsymbol{\pi}'_{ij} = (\boldsymbol{\pi}_{ij}, 1)$$

$$\hat{y}_i = \boldsymbol{\eta}_{0:K-1}^\top \mathrm{E}_q[\bar{Z}_i] + \eta_K = \boldsymbol{\eta}_{0:K}^\top \frac{1}{L_i}\sum_{j=1}^{L_i}\boldsymbol{\pi}'_{ij} \qquad (16)$$

モデル予測性能の評価指標をラベルの真値と予測値との間の平均二乗誤差 (Mean Squared Error : MSE) を用いる：

$$MSE = \frac{1}{D}\sum_{i=1}^{D}(y_i - \hat{y}_i)^2.$$

すべての実験は MSE の変化率が 0.1%以下になるときに収束と判断され終了する．

4.3 実験結果

本実験では，学習データを学習するときに，5 回おきに得られた $\boldsymbol{\pi}$ と $\boldsymbol{\eta}$ を用いて MSE を計算した．各条件における TopicVec-Reg の性能を同じ条件の TopicVec+LR と比較した．

3 つのデータセットにおける TopicVec-Reg と TopicVec-LR を用いたそれぞれ 15 通りの実験に対して MSE を算出し，さらに同じ条件における両手法の予測値に対して

Wilcoxon 符号付き順位検定を行った．実験結果として得られた MSE と p 値を表 2 に示す．太字で TopicVec-Reg が TopicVec+LR より良い予測精度を示す．太字の p 値が 5%を下回っており有意な差を示す．

表 2: 実験結果

K		set1	set2	set3
5	TopicVec-Reg	**0.016637**	**0.017398**	0.016179
	TopicVec+LR	0.016867	0.017730	0.016107
	p-value	**0**	**0**	0
10	TopicVec-Reg	0.016343	**0.015873**	**0.015685**
	TopicVec+LR	0.016220	0.016879	0.015839
	p-value	0.870397	**0.000693**	0.086139
15	TopicVec-Reg	**0.015873**	**0.015297**	**0.014299**
	TopicVec+LR	0.016273	0.016803	0.015616
	p-value	**0.020224**	**0.038315**	0.1617699
20	TopicVec-Reg	**0.013736**	**0.015790**	**0.013564**
	TopicVec+LR	0.016453	0.016785	0.014910
	p-value	**0.000565**	0.486916	**0.016891**
25	TopicVec-Reg	**0.013447**	**0.014533**	**0.012280**
	TopicVec+LR	0.015963	0.016797	0.014726
	p-value	**0.044846**	0.492524	**0.001298**

5 むすび

本稿では，金融記事から企業の株価上昇率を予測する課題に取り組み，トピック埋め込みモデル TopicVec に回帰機能を付与することで，トピック埋め込み回帰モデル TopicVec-Reg を提案した．

TopicVec-Reg では潜在トピックと回帰パラメータを同時に学習することができ，実験により潜在トピックを推定した後に回帰パラメータを推定するモデルより有意な改善を得た．

本稿ではクローズドテストにより評価を行ったが，オープンテストは今後の課題である．また，潜在トピックだけではなく，単語埋め込みを活用できるように説明変数を増やして回帰性能のさらなる改善を目指したい．

謝辞

本研究の一部は科学研究費補助金基盤研究 (B)（15H02703）の援助による．

参考文献

[1] David M Blei, Andrew Y Ng, and Michael I Jordan, ”Latent Dirichlet Allocation”, *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022 (2003)

[2] Dieng, Adji B., Francisco JR Ruiz, and David M. Blei. ”Topic modeling in embedding spaces.” *arXiv preprint arXiv:1907.04907* (2019)

[3] Mikolov, Tomas, et al. ”Distributed representations of words and phrases and their compositionality.” *Advances in neural information processing systems.* (2013)

[4] David M Blei and Jon D. McAuliffe, ”Supervised topic models”, *Advances in neural information processing systems*, pp. 121–128 (2008)

[5] Das, Rajarshi, Manzil Zaheer, and Chris Dyer. ”Gaussian lda for topic models with word embeddings.” *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (2015)

[6] Li, Shaohua, et al. ”Generative topic embedding: a continuous representation of documents (extended version with proofs).” *arXiv preprint arXiv:1606.02979* (2016)

[7] Li, Shaohua, Jun Zhu, and Chunyan Miao. ”A generative word embedding model and its low rank positive semidefinite solution.” *arXiv preprint arXiv:1508.03826* (2015)

[8] 佐藤敏紀, 橋本泰一, 奥村学, ”単語分かち書き辞書 mecab-ipadic-NEologd の実装と情報検索における効果的な使用方法の検討”, 言語処理学会第 23 回年次大会 (NLP2017), NLP2017-B6-1 (2017)

[9] 佐藤敏紀, 橋本泰一, 奥村学, ”単語分かち書き用辞書生成システム NEologd の運用 – 文書分類を例にして –”, 自然言語処理研究会研究報告, NL-229-15 (2016)

[10] Toshinori, Sato, ”Neologism dictionary based on the language resources on the Web for Mecab”(2015)

ニューステキストを用いたESGファクター運用

ESG factor investing using news text data

秋山祥伍 [1*]　江口潤一 [2]　鈴木智也 [1,2]
Shogo Akiyama[1]　Junichi Eguchi[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] 大和証券資信託委託株式会社
[2] Daiwa Asset Management Co.Ltd.

Abstract: ESG investment is getting popular as an investing method that evaluates environmental, social and governance (ESG) efforts of each company. However, although the disclosure of information on ESG is already common, it is difficult to objectively evaluate ESG efforts of companies due to the ambiguity in the definition of ESG words: environment, society, and governance. For this reason, we applied Word2Vec to extract similar words to each ESG word, and evaluated ESG efforts of companies from the viewpoint of these extracted words. As a result, we confirmed that the companies with higher ESG score can make a more profitable portfolio than those with lower ESG score, and this ESG score can be useful for factor investing.

1　はじめに

ESG 投資は企業の環境（Environment），社会（Social），ガバナンス（Governance）への取り組みを評価した投資である．近年，企業の社会的責任（CSR）が重要視され，特に ESG は重要なキーワードである．

企業の ESG 文書開示数により ESG への積極性を評価した研究 [1] がある一方で，ESG というキーワードの曖昧さにより，各企業の開示文書の内容から ESG への取り組みを評価することは難しい．文書を各トピックに割り当てる手法としてトピックモデルがある．トピックモデルの応用例として Word2Vec[2] などの単語の分散表現を用いたトピックモデルの構築 [3] や，ファンドの割り当てへの応用 [4] などがある．

そこで本研究では，ニューステキストの ESG トピックワードに着目した企業の ESG スコアの構築を行う．ESG のトピックワードの抽出には Word2Vec を使用し，抽出した単語を用いてニューステキストの評価に使用する．また，その結果から ESG ファクターを構成し，これが既存の代表的なファクターに対して補完的に機能するかを検討する．

*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: 19nm401h@vc.ibaraki.ac.jp

2　ESG 単語の抽出

2.1　単語の分散表現

単語の記号から類似度や関連性を計算するためには任意のベクトルで表現する必要があり，これを分散表現という．特に単語の分散表現を獲得する代表的な手法に Word2Vec[2] がある．Word2Vec はニューラル言語モデルであり，CBOW モデルと Skip-gram モデルの 2 つのモデルが提案されている．ネットワーク構造は Encoder と Decoder から構成され，Encoder は単語の記号から分散表現への変換，Decoder は分散表現から記号への変換を行う．Word2Vec から得られた単語の分散表現では，図 1 のように似た単語同士は近いベクトルに表され，それらは一種のトピックの役割を果たす．したがって，本研究ではその特性を利用して分散表現から「環境」，「社会」，「ガバナンス」の周辺単語を取り出すことで，ESG スコアの評価に使用する．

Word2Vec の問題として類義語だけでなく対義語も近いベクトルに変換されてしまう（例 : 好き・嫌い，高い・安い）．これは，ESG にネガティブなニュースに対しても高く評価しかねない．この問題を解決するために Word2Vec の学習においてはマーケットに対してポジティブなニュースのみを使用することでネガティブ単語を除去する．ポジティブなニュースを抽出する方法として QUICK ニュース解析サービス [5] が提供しているニューススコアを利用する．

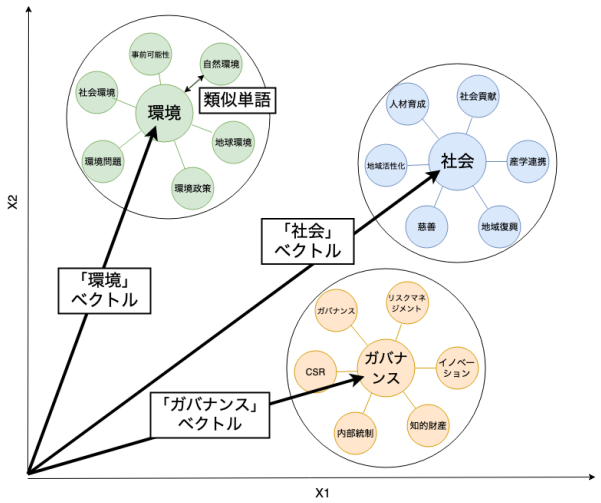


図 1: Word2Vec の概要

2.2 結果

Word2Vec の学習には 2011 年 1 月 1 日 ~2012 年 12 月 31 日の QUCIK ニュース解析サービスから配信されたニュースデータを使用する．Word2Vec による「環境」,「社会」,「ガバナンス」と cos 類似度が大きい上位 10 単語を表 1 に示す．

「環境」との cos 類似度が最も高い単語が「加速度」となった理由は,「加速度的増加」などのニュース特有の言い回しがあるためである．また,「社会」との cos 類似度が最も高い単語が「炭素」となった理由は,「低炭素社会」という単語がテキスト中に含まれるためである．

表 1: 各 ESG に類似する上位 10 単語

環境	社会	ガバナンス
加速度	炭素	委譲
永続	我が国	実効
リユース	CSR	総力
クラウドコンピューティング	事業	統制
保全	計算	戦略
親和	資本	選抜
国連	当期	疎通
低減	利益	法務
テレワーク	証券	一元化
カーボンフットプリント	評価	分業

3 ESG 評価

3.1 ESG スコア

Word2Vec を用いて「環境」,「社会」,「ガバナンス」に類似する上位 100 単語を取り出し，各企業の ESG スコアを計算する．例として「環境」スコアの計算方法を示す．ニュース i に含まれる Word2Vec から抽出した「環境」類似単語数を W_i,「環境」単語のベクトルを $\boldsymbol{w}_E$, 類似単語のベクトルを $\boldsymbol{w}_j$, 類似単語を含むニュース数を N とすると，あるニュースの「環境」スコア $Score_E$ を式 (1) のように定義する．

$$Score_E = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{W_i} cos(\boldsymbol{w}_E, \boldsymbol{w}_j) \quad (1)$$

ここで $cos(\boldsymbol{v}_a, \boldsymbol{v}_b)$ は単語 a と b のベクトル $\boldsymbol{v}_a, \boldsymbol{v}_b$ の cos 類似度を表す．

$$cos(\boldsymbol{v}_a, \boldsymbol{v}_b) = \frac{\boldsymbol{v}_a \cdot \boldsymbol{v}_b}{|\boldsymbol{v}_a||\boldsymbol{v}_b|} \quad (2)$$

なお,「社会」,「ガバナンス」も同様に計算する．

ESG は企業の長期的取り組みであることから，企業の ESG スコアは 2011 年 1 月 1 日から投資する月までの式 (1) の合計で計算し，スコアは月末に更新する．

3.2 ESG ポートフォリオ

上記で計算した企業の ESG スコアを上位 30%(High), 中位 40% (Medium), 下位 30% (Low) に分け，それぞれのパフォーマンスを計測する．なお，企業のニュース配信数によるサイズ効果をなくすため，ニュースが月に一度も配信されない銘柄はあらかじめ除去した．

期間は 2013 年 1 月 1 日 ~2018 年 12 月 31 日までの月足データを使用し，上記の ESG ポートフォリオについて累積収益率 $M_1(t)$ を計算する．

$$M_1(t) = M_1(1) + M_1(1)\prod_{\tau=1}^{t}\{1 + r(\tau)\} \quad (3)$$

$$r(t) = \frac{P(t) - P(t-1)}{P(t-1)} \quad (4)$$

ここで，$P(t)$ は時刻 t における株価, $r(t)$ は収益率であり，初期資産は $M_1(1) = 1$ とする．

得られた結果を図 2 に示す．ESG スコアが高い High が低い Low よりも高い収益率を得ており，Word2Vec で得られた ESG 単語が十分に機能してと示唆される．

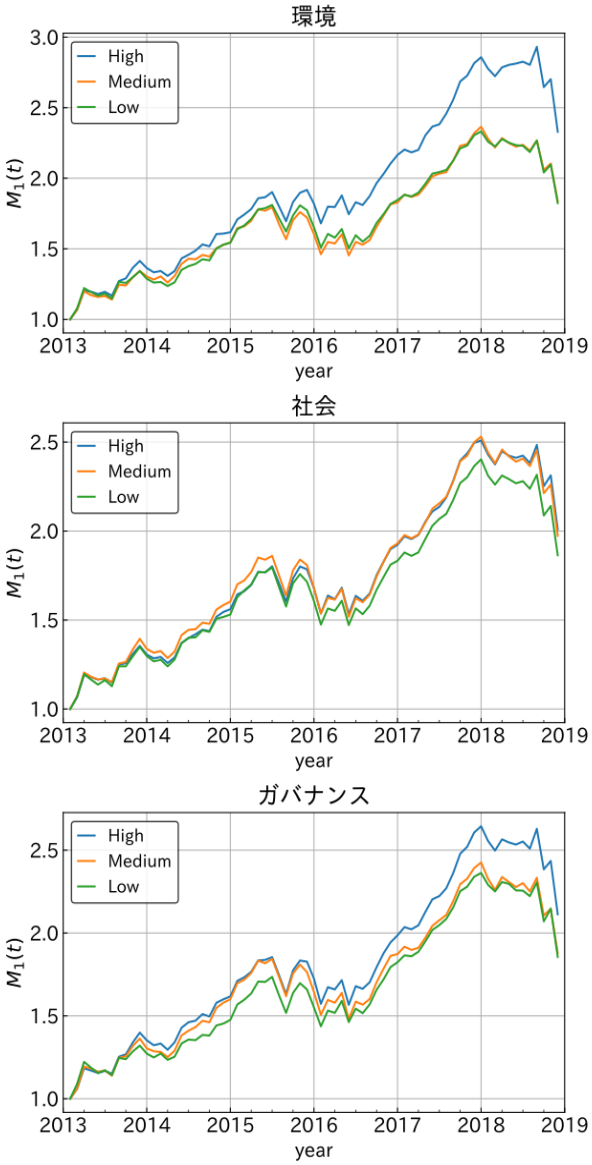


図 2: ESG ポートフォリオの累積収益率 $M_1(t)$

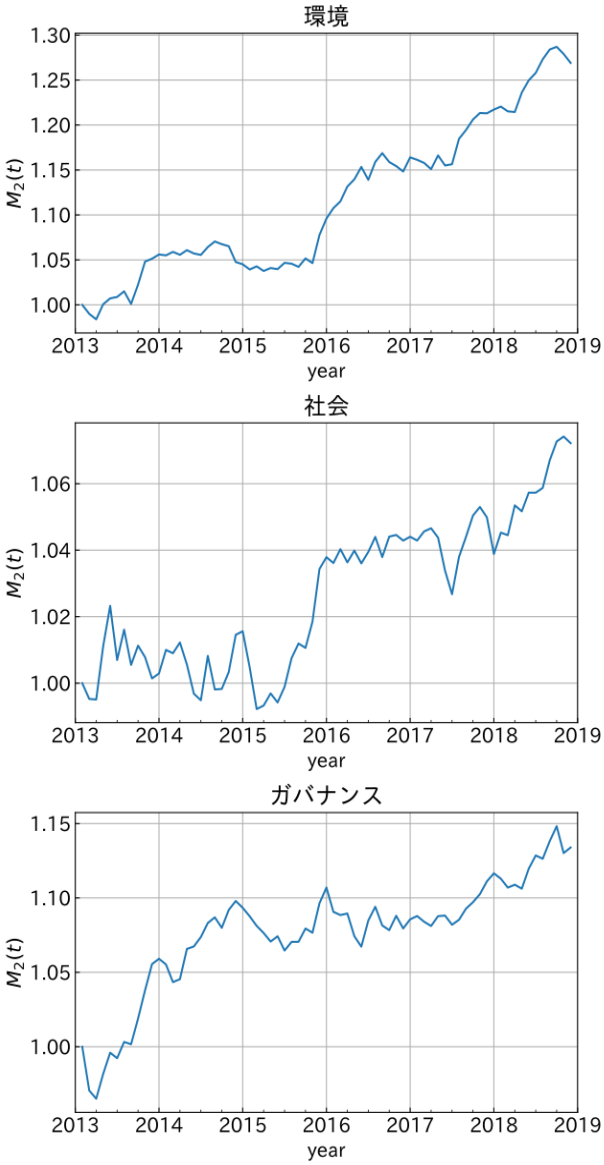


図 3: ESG ファクターの累積収益率 $M_2(t)$

次に，High から Low のリターンを引くことにより，ESG ファクターとして累積収益率 $M_2(t)$ を計測した．

$$M_2(t) = M_2(1) + M_2(1)\prod_{\tau=1}^{t}\{1+s(\tau)\} \quad (5)$$

$$s(t) = \frac{1}{N_H}\sum_{i\in High} r_i(t) - \frac{1}{N_L}\sum_{j\in Low} r_j(t) \quad (6)$$

$r_i(t)$ は時刻 t におけるポートフォリオ i の収益率，N_H は上位ポートフォリオの銘柄数，N_L は下位ポートフォリオの銘柄数，$s(t)$ は各 ESG ファクターのリターンである．初期資産は $M_2(1) = 1$ とする，

ここで，得られた $s(t)$ の平均，標準偏差，平均/標準偏差を表 3 に，累積収益率 $M_2(t)$ を図 3 に示す．これらの結果から，ESG の各ファクターにおいて正の収益を期待でき，特に「環境」と「ガバナンス」に関して安定的なリターンを確認できる．

3.3 マルチファクターモデル

前章で得られた各企業の ESG スコアの有効性を 3 ファクターモデル [6] により検証する．3 ファクターモデルは次式の通りである．

$$r_{it} - r_{ft} = \beta_i + \beta_i^M(r_{Mt} - r_{ft}) + \beta_i^{SMB}SMB_t + \beta_i^{HML}HML_t + \epsilon_{it} \quad (7)$$

表 2: ファクターの相関係数

	マーケット	サイズ	バリュー	環境	社会	ガバナンス
マーケット	1.00	-0.01	-0.14	-0.18	-0.38	-0.15
サイズ	-0.01	1.00	-0.46	-0.05	-0.14	-0.31
バリュー	-0.14	-0.46	1.00	-0.23	0.21	0.36
環境	-0.18	-0.05	-0.23	1.00	0.29	0.37
社会	-0.38	-0.14	0.22	0.29	1.00	0.32
ガバナンス	-0.15	-0.31	0.36	0.37	0.32	1.00

表 3: ESG ファクターのリターン

	平均 (%)	標準偏差 (%)	平均/標準偏差
環境	0.31	0.98	0.32
社会	0.09	0.67	0.13
ガバナンス	0.18	0.92	0.20

表 4: ESG ファクターの 3 ファクターモデルによる偏回帰係数

	環境		社会		ガバナンス	
	係数	t 値	係数	t 値	係数	t 値
β_i	0.36	3.16***	0.15	1.91	0.24	2.32*
β_i^M	-0.06	-1.99*	-0.07	-3.24***	-0.03	-1.01
β_i^{SMB}	-0.10	-1.75	-0.03	-0.73	-0.08	-1.54
β_i^{HML}	-0.14	-2.81***	0.03	0.96	0.09	2.03*

*$p < 0.05$, ***$p < 0.001$

ここで，$r_{Mt} - r_{it}$ はマーケットファクター，SMB_t はサイズファクター，HML_t はバリューファクターであり，$\beta_i^M, \beta_i^{SMB}, \beta_i^{HML}$ はそれぞれの偏回帰係数を表す．

ESG ファクターの有効性を検証するために，式 (7) の r_{it} を ESG ファクターのリターン $s(t)$ として重回帰分析を行う．この分析結果を表 4 に示す．「環境」と「ガバナンス」に関しては，切片が有意に正となり，3 ファクターでは説明できないリターンが存在することが示唆される．また，3 ファクターと $s(t)$ との相関係数を表 2 に示す．この結果として，作成した ESG ファクターは他のファクターとの相関が小さいため，独自のファクターとして機能していることが分かる．

4 まとめ

本研究では Word2Vec によりニューステキストから抽出した ESG トピックワードを用いて各企業の ESG スコアを評価し，その評価に基づいて構成した ESG ファクターを分析した．その結果，「環境」および「ガバナンス」ファクターが既存の代表的なファクターでは説明できないリターンを有することが確認できた．

今後の課題として，「社会」ファクターのパフォーマンスの向上や単語だけでなく文書構造を考慮した ESG スコアの評価などが挙げられる．

なお本研究の一部は，文化省科研費基礎研究 (C) (No.16K00320) の助成により行われました．

参考文献

[1] 湯山智教, 白須洋子, 森平総一郎. ESG 開示スコアとパフォーマンス, 証券アナリストジャーナル, Vol. 57, No. 10, pp. 72–83, 2019

[2] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient Estimation of Word Representations in Vector Space, *International Conference on Learning Representations*, pp. 1–12, 2013.

[3] Rajarshi Das, Manzil Zaheer, and Chris Dyer. Gaussian LDA for Topic Models with Word Embeddings, *In Proceedings of the 53nd Annual Meeting of the Association for Computational Linguistics*, pp. 795–804, 2015.

[4] Masanori Hirano, Hiroki Sakaji, Shoko Kimura, Kiyoshi Izumi, Hiroyasu Matsushima, Shintaro Nagao, and Atsuo Kato. Related Stocks Selection with Data Collaboration Using Text Mining, *Information*, Vol. 10, No. 3, pp. 1–17, 2019

[5] 株式会社 QUICK, QUICK ニュース解析サービス, https://nwfs.newsdolphin.jp/ja, (参照 2019-02-23)

[6] Eugen F.Fama and Kenneth R. French. Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, Vol. 33, No. 1, pp. 3–56, 1993

グラフエンベディングを活用した潜在取引関係予測

Latent Transaction Prediction Using Graph Embedding

藤塚 理史[1]　工藤 剛[2]

Masashi Fujitsuka[1], Tsuyoshi Kudo[2]

[1] 株式会社日本総合研究所
[1] The Japan Research Institute, Limited
[2] 株式会社三井住友銀行
[2] Sumitomo Mitsui Banking Corporation

Abstract: Analysis of transaction data between corporations has strong future growth potential in finance since financial institutions have the large amount of the data. However, since it is difficult for each institution to get such data except the main customers' one, the data is partial and the application would be also limited. We set a similar environment artificially from the only observed data by removing some observed transaction links, and evaluate if we could predict the removed links. We show that graph embedding could lead to a solution of this problem.

１．はじめに

金融機関の主要な業務として顧客である企業の与信がある．この与信業務は，各企業の財務情報に基づいてしばしば行われるが，企業ごとのそれぞれの情報だけでなく，その企業がどのような企業と取引があるか，さらにその取引先企業はどのような企業と取引があるかといった取引関係データを活用することにより，従来の与信業務を高度化が出来る可能性がある．例えば，現状の与信判断における精度向上，システムによる自動化，また与信判断材料が不足している企業群に対する与信業務の適用拡大といったことなどが考えられる．取引関係と企業成長に関しても研究が行われており[1]，取引関係データの活用は大きく期待される．

しかし，取引関係データを活用する際の大きな課題として，全ての企業間の取引関係を把握することが出来ないことが挙げられる．例えば，1 つの金融機関では，その機関をメインバンクとしない企業の主要な取引関係を把握することは出来ないため，得られている取引関係のデータ量は多い一方で，多くの取引関係が未観測で欠損した状態になっている．そのため，取引関係データの活用範囲は限定的なものになってしまう．

上記のような課題に対して，観測出来ていないが実際には存在している取引関係（以下、潜在取引関係と記す）を予測することが出来れば，取引関係データの活用の幅を大きく広げることが可能となる．

そこで本研究では，潜在取引関係を擬似的に設定することにより，その潜在関係を予測する問題を考える．そして，そのような予測問題に対して，グラフエンベディング[2]と呼ばれる関係性データ（グラフデータ）から特徴を抽出する手法が有効であることを示す．

２．課題に対するアプローチ方法

問題設定

ここでのタスクは，潜在取引関係の予測を行うことであるが，潜在取引関係に当たるデータは入手が困難なため，実際にそのような関係を予測して評価を行うことは出来ない．そこで，次のように擬似的に潜在取引関係を設定することで，その関係を予測する問題に置き換える：

- ある企業に着目し，その企業が持つ取引関係をある割合（マスク率）でマスクする
- マスクされた取引関係を，潜在取引関係と見なし，それらの予測問題とする

これにより，未観測の取引関係を予測する問題を，現状保有するデータにおけるマスクされた取引関係を予測する問題と見なして考える．実際の課題と照らし合わせた場合，他の金融機関をメインバンクとするような企業ならば，マスク率が高い場合に相当すると考えられる．

グラフ分析としてのタスク

今回扱う取引関係データは，いわゆるグラフデータと呼ばれるデータ形式に対応する．グラフデータとは，ノードの集合 V と，リンクの集合 E の組 (V,E) からなる．各企業がノード，企業間取引がリンクに対応するグラフと見なすことが出来る．特に，今回の取引関係の予測は，グラフ分析におけるリンク予測の問題に対応する．

3．グラフ分析手法

3−1．リンク予測における指標

リンク予測とは，あるグラフデータが与えられた時に，それを手がかりにして，まだ知られていないリンクが存在するどうかを予測する問題である[3]．ノード u の隣接ノード集合を $\Gamma(u)$，隣接ノード数を $|\Gamma(u)|$ とする．ノードペア (u,v) に対して，リンクが存在するかどうかを表す指標として，下記のようなものが提案されている：

- 共通隣接ノード指標[4]

$$|\Gamma(u) \cap \Gamma(v)|$$

2 つのノードの間における共通の隣接ノード数を表したものである．つまり，共通の隣接ノードが多いノード同士はつながる可能性が高いことを表す．

- Jaccard 係数[5,6]

$$\frac{|\Gamma(u) \cap \Gamma(v)|}{|\Gamma(u) \cup \Gamma(v)|}$$

共通隣接ノード指標を，ノード u とノード v の隣接ノード全体で正規化したものである．つまり，2 つのノードの隣接ノードの全体の内，大半が重なるなら，つながる可能性が高いことを表す．

- Adamic/Adar 指標[7]

$$\sum_{w \in \Gamma(u) \cap \Gamma(v)} \frac{1}{\log |\Gamma(w)|}$$

共通隣接ノードの各ノードごとに異なる重みを設定し，足し合わせた指標である．例えば，どのノードともつながるようなノードとの共通リンクは，めったにリンクを持たないノードのリンクに比べて情報は多くないと考えられる．この指標は，そのような重みを考慮して設計された指標となっている．

- 優先的選択指標[8]

$$|\Gamma(u)| \cdot |\Gamma(v)|$$

この指標は，共通しているかどうかは関係なく，そもそも隣接ノード数が多ければ，リンクが存在しやすいとする指標である．例えば，どのノードともつながりやすいノードとは，そもそもつながる可能性が高いことを表す．

3−2．グラフエンベディング

上記で挙げたリンク予測の指標は，基本的に各ノードの隣接ノードの情報しか用いていない．一方で，今回の課題においては，隣接ノードの多くがマスクされて見えていないような状況であるため，このような指標は有効ではないように考えられる．そのため，隣接ノードの情報だけでなく，さらにそれより先の隣接関係性まで考慮することが必要である．そこで，そのような高次の隣接関係性まで考慮できるグラフエンベディングと呼ばれる手法をここでは考える．

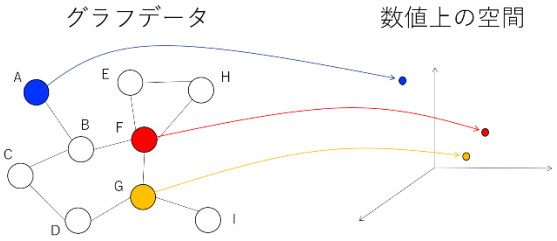


図 1：グラフエンベディングのイメージ

グラフエンベディングとは，グラフ上の性質を反映させるように，ノードごとに数値ベクトルを割り当てるものである．保ちたい性質に依存して，様々な手法が存在する．代表的な例としては，隣接ノードの関係性を保つために，以下のような最適化問題を解くことで，d 次元のエンベディング $f: V \to \mathbb{R}^d$ を求める Laplacian Eigenmaps[9]がある：

$$\min_f \sum_{(u,v)\in E} |f(u) - f(v)|^2.$$

近年，隣接ノード関係だけでなく，より高次の隣接関係性を保つようなグラフエンベディングの手法が複数提案されている[2,10-16]．ここでは特に，計算コストが少なく，柔軟にグラフの関係性を捉えることが可能な node2vec[2]と呼ばれるグラフエンベディング手法を用いる．node2vec は，ランダムウォークによるグラフ上の探索と，その探索された範囲における（広い意味での）隣接関係を保つようなエンベディングの学習からなる：

I.　ランダムウォークによるグラフ上の探索

グラフ上での性質を捉えるために，グラフ上でランダムウォークを行い，グラフデータをノードの系列データの集合に変換する．これにより，グラフ上での隣接関係性は，系列データにおける隣接関係性として置き換えられる．

この系列データにおける隣接関係性は，ランダムウォークに大きく依存する．node2vec では，ノード位置 u_{t-1} から u_t に移動した時，元の位置 u_{t-1} からの距離に基づいて，次に移動可能なノード x に対して，次のようなバイアスを設定する：

$$\alpha_{pq}(t,x)=\begin{cases}\dfrac{1}{p} & (if\ \ d(u_{t-1},x)=0)\\ 1 & (if\ \ d(u_{t-1},x)=1)\\ \dfrac{1}{q} & (if\ \ d(u_{t-1},x)=2)\end{cases}$$

ここで，$d(u,v)$ はノード u と v の距離を表す．パラメータ (p,q) がバイアスパラメータであり，例えば，p が q（かつ1）に比べて小さい場合，元いた位置に戻りやすくなる．逆に，大きい場合は，遠方に行きやすくなる．

このような探索により，単純な隣接ノードの関係性でなく，より広い意味でのグラフ上での隣接関係性を捉えられると期待できる．今回は，近傍，遠方に特に偏りがないように探索したいため，$p=q=1$ に設定する．

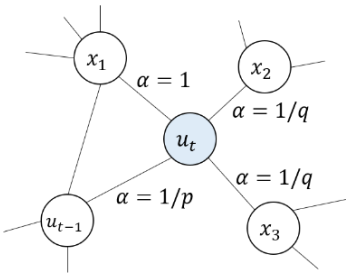


図 2：バイアスランダムウォーク

II.　探索範囲におけるエンベディングの学習

系列データからエンベディングを学習する手法として word2vec[17]を用いる．word2vec では，似たような文脈で出現する単語の表現は似ているとする分布仮説に基づき，単語列の前後で現れる単語を予測するモデル（Skip-gram モデル）を学習することで，エンベディングを得る[1]．この時の前後のコンテキストはウィンドウサイズを設定することにより柔軟に設定が可能である．

今回の場合は，ランダムウォークで得られた隣接関係性に基づいて，似たような関係性を持つノードが似たような表現になるようにグラフエンベディングが学習されることになる．これにより，3-1 に記載した隣接ノード関係だけを考慮した指標では捉えられないようなノード間の関係性が捉えられると期待できる．

４．実験

４−１．実験設定

ここでは，ある金融機関の 1 年間の企業間取引関係データを用いる．年間で送金、もしくは着金がある一定金額以上ある場合に、2 つの企業間に取引があるとして重みなし無向リンクを引く．このようにして作成した企業間取引関係グラフを今回の実験対象とする．企業数約 100 万社，約 1,000 万取引関係を含むグラフとなっている．

学習データ

企業を 1 社ランダムに選択し，その企業が持つ取引関係をあるマスク率でマスクする．ここでマスクされた取引関係が，正解となるテストデータに対応する．それに加えて，教師あり機械学習用の訓練データとして，ランダムに 10,000 個の取引関係を選択し，それらもマスクする．このようにして，マスクされたグラフデータが学習データに対応する．

評価データ

評価方法としては，上述でランダムに選ばれた 1 社に対して，正例（リンクが存在するノードペア）と負例（リンクが存在しないノードペア）の集合を与えた場合に，正例をいかに多く当てることが出来るかどうかの 2 値分類問題として評価を行う．正例は，上述のマスクした取引関係である．負例は，取引関係を持たない企業からランダムにサンプリングしたものに加えて，取引関係を持たない取引社数が多い企業からサンプリングしたものを同じ割合で用意して，それらを負例とする．

今回の対象企業は，図 3 のように取引社数が多い企業が指数関数的に減少するような分布となっている．そのため，ランダムにサンプリングすると，取引社数が少ない企業が多く抽出される．一方で，取引社数が多い企業とはリンクがつながりやすい傾向

[1] word2vec は，Skip-gram と Continuous Bag of Words（CBOW）の 2 つの定式化があるが，node2vec では前者を用いたモデルが採用されている．

があるため，正例の多くは取引社数が多い企業となりやすい．

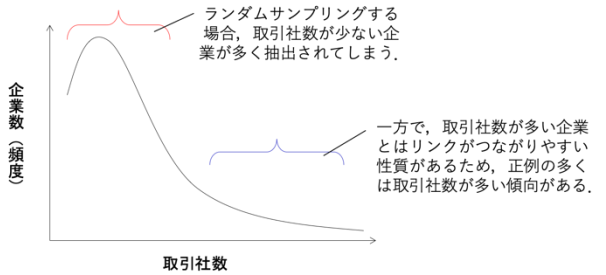


図 3：取引社数ごとの頻度分布（次数分布）

つまり，ランダムサンプリングを行うと，取引社数の大小によって，正例と負例が分かれる傾向となってしまう[2]．今回は，そのような傾向の違いを捉えられるかどうかを検証したいわけでないため，取引社数が多い企業から抽出した負例を加えたものを評価データとする．

評価方法

上述した評価は，1 つの企業に対する評価であり，サンプリングされた企業によって結果は大きく依存する．そのため，上述の評価をランダムにサンプリングした 20 社に対して独立に行い，その平均 ROC-AUC スコアで評価を行う．

具体的な評価としては，取引社数ごとに，100-300 社，300-600 社，600-1000 社の 3 つのカテゴリに分けて，それぞれで実施する．そして，それぞれの評価において，取引関係のマスク率を，20%，50%，80%とそれぞれ変えて評価を行う．

グラフエンベディング（node2vec）の学習

実装は，Stanford Network Analysis Project（SNAP）による C++実装を用いる[3]．上述したように，重みなし無向グラフとし，ランダムウォークのバイアスパラメータは，$p = q = 1$．探索は，各ノードを始点とした 80 長さのウォークを 10 セット実施．エンベディングのベクトル次元は 128，系列データの隣接ウィンドウサイズは 10 として学習を実施した．

4-2．比較手法

グラフエンベディング手法 node2vec を用いた場合と，3-1 に記載した 4 つの指標（以下，ベースライン指標と記す）との比較を行う．node2vec に関しては，2 つのノードから得られるそれぞれのエンベディングベクトル $f(u), f(v)$ から，cos 類似度，

$$\frac{f(u) \cdot f(v)}{|f(u)|\ |f(v)|}$$

を算出することで指標として用いる．

また，教師あり機械学習の評価として，

- ベースライン指標を特徴量としたモデル
- ノードペアそれぞれの node2vec で得られるエンベディングの要素積を取り，各次元をそれぞれ特徴としたモデル[4]

の 2 つを評価する．教師あり機械学習モデルとしては，勾配ブースティングモデルを使用する．学習用の正例データは，4-1 の学習データ作成の際に除去した 10,000 取引関係，負例データは，取引関係がない企業ペアからランダムにサンプリングした 10,000 取引関係を用いる．

4-3．実験結果と考察

表 1 に実験結果を記載した．ベースライン指標に比べて，高い精度で予測出来ていることが分かる[5]．特筆すべきは，マスク率を上げた場合の，精度の減少度合いである．ベースライン指標では，およそ 0.05 程度減少している一方で，グラフエンベディングを用いた場合，その半分程度である．これは，保有している取引データが部分的なものである場合でも，グラフエンベディングを用いることで取引関係を予測出来ることを示唆している[6]．

ベースライン指標は，隣接ノードの関係性しか見ていないため，マスク率が高くなるに従って，その予測精度は急激に減少してしまっていると考えられる．一方で，グラフエンベディング（node2vec）は，マスク率が高い状況でも，その先の隣接関係性まで捉えていると考えられるため，減少度合いが少なく，有効に機能すると考えられる．

[2] この場合，リンク数が多いほど高い確率とする優先的選択指標が最も有効となる．
[3] https://github.com/snap-stanford/snap/tree/master/examples/node2vec
[4] node2vec を用いたリンク予測の特徴量としては，要素積を取る以外の方法もあるが，[2]におけるリンク予測の実験で要素積による特徴が最も精度が高い結果であったことから，今回は要素積を選択する．
[5] 優先的選択指標に関しては，相手先の企業の取引数にしか依存しないため，マスク率よる影響はほとんど受けない結果となる．
[6] 今回の結果は，ランダムに選択された 20 社の平均精度である．1 つ 1 つの結果で見た場合，ベースライン指標の方が高いケースもいくつか存在する．

取引社数範囲	マスク率	教師なし学習					教師あり学習	
		ベースライン指標				グラフエンベディング cos 類似度	ベースライン指標	グラフエンベディング
		共通隣接ノード指標	Jaccard 係数	Adamic/Adar 指標	優先的選択指標			
100-300 社	20%	0.662	0.679	0.671	0.565	**0.796**	0.684	**0.843**
	50%	0.657	0.663	0.665	0.563	**0.794**	0.666	**0.844**
	80%	0.613	0.601	0.620	0.565	**0.773**	0.617	**0.815**
300-600 社	20%	0.657	0.695	0.673	0.523	**0.782**	0.690	**0.838**
	50%	0.639	0.673	0.651	0.526	**0.770**	0.656	**0.824**
	80%	0.606	0.622	0.614	0.525	**0.760**	0.594	**0.816**
600-1000 社	20%	0.644	0.671	0.652	0.518	**0.765**	0.679	**0.802**
	50%	0.635	0.670	0.642	0.514	**0.750**	0.655	**0.798**
	80%	0.588	0.616	0.592	0.512	**0.742**	0.577	**0.783**

表 1：実験結果（20 社の平均 ROC-AUC スコア）

残された課題

今回の評価方法は，着目した企業に対して，取引関係をあるマスク率でマスクするというような特定の状況を考えたが，現実的には，ある金融機関をメインバンクとしない企業は複数存在するため，マスクの取り方をより複雑にした設定における評価が必要になると考えられる. また, リンク予測の性質上, 取引社数が少なすぎると, うまく機能しない. また, 多すぎる場合も問題が困難になると考えられるため, 適用可能範囲を明確にすることも求められる.

また，今回は企業ごとのつながりの情報しか使っていないが, 企業が持つ属性情報も活用することで, より高度な予測につなげられると考えられる．例えば，今回の分析では，口座を持たない企業，つまりグラフ内にノードが存在していないものは自動的に対象外とするが，属性情報用いることで企業間の類似度の算出が可能となり，全く関係性を持たないような企業に対してもアプローチが可能となるといった方向性も期待出来る.

５．結論

本研究では，取引関係データが部分的にしか得られないという課題に対して，観測出来ていないが実際には存在する取引関係（潜在取引関係）を保有している取引関係データだけから疑似的に設定することにより，それらを予測する問題として定式化し，その問題に対し，グラフエンベディングと呼ばれる手法が有効であることを示した．これにより，現状保有しているデータ量だけでは見えない部分の補完が可能となることから，取引関係データの活用の幅を大きく広げることが出来ると考えられる.

参考文献

[1] みずほ情報総研株式会社，平成２８年度 中小企業の成長に向けて中長期的に 取り組むべき施策の検討に向けた 我が国中小企業の成長過程分析に係る委託調査, 経済産業省平成 28 年度委託調査報告書, 2018.

[2] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2016.

[3] 鹿島久嗣，ネットワーク構造予測，人工知能学会誌 Vol.22 No.3, 2007.

[4] Newman, M. E. J.: Clustering and preferential attachment in growing networks, Physical Review Letters E, Vol. 64(025102), (2001)

[5] Baeza-Yates, R. A. and RibeiroNeto, B. A.: Modern Information Retrieval, A 8 Press / Addison-Wesley (1999)

[6] Liben-Nowell, D. and Kleinberg, J.: The Link Prediction Problem for Social Networks, in Proceedings of the Twelfth International Conference on Information and Knowledge Management (CIKM), pp. 556–559 (2004)

[7] Adamic, L. A. and Adar, E.: Friends and neighbors on the Web, Social Networks, Vol. 25, No. 2, pp. 211– 230 (2003)

[8] Barab´asi, A. L., Jeong, J., N´eda., Z., Ravasz, E., Shubert, A., and Vicsek, T.: Evolution of the social network of scientific collaborations, Physica A, Vol. 311, No. 3-4, pp. 590–614 (2002)

[9] M. Belkin and P. Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In NIPS, 2002.

[１０] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 701–710. ACM, 2014.

[１１] Ahmed, N. Shervashidze, S. Narayanamurthy, V. Josifovski, and A.J. Smola. Distributed large-scale natural graph factorization. In WWW, 2013.

[１２] S. Cao, W. Lu, and Q. Xu. Grarep: Learning graph representations with global structural information. In KDD, 2015.

[１３] M. Ou, P. Cui, J. Pei, Z. Zhang, and W. Zhu. Asymmetric transitivity preserving graph embedding. In KDD, 2016.

[１４] J. Tang, M. Qu, M. Wang, M. Zhang, J. Yan, and Q. Mei. Line: Large-scale information network embedding. In WWW, 2015.

[１５] S. Cao, W. Lu, and Q. Xu. Deep neural networks for learning graph representations. In AAAI, 2016

[１６] D. Wang, P. Cui, and W. Zhu. Structural deep network embedding. In KDD, 2016.

[１７] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In NIPS, 2013.

日本株式市場における daytime と overnight return の間の VC 相関解析

Volatility Constrained correlation on Japanese stock market

落合　友四郎 [1*]　ホセ・ナチェル [2 †]
Tomoshiro Ochiai[1]　Jose C. Nacher[2]

[1] 大妻女子大学社会情報学部
[1] Faculty of Social Information Studies, Otsuma Women's University
[2] 東邦大学理学部情報科学科
[2] Department of Information Science, Faculty of Science, Toho University

Abstract: Although econophysics techniques have been widely applied to investigate stock markets, the analysis of non-trading or night periods has received less attention. Here, we investigate the correlation between overnight and daytime return (correlation ND) and the correlation between daytime return and following overnight return (correlation DF). A standard correlation analysis reveals a weak negative correlation between overnight and daytime return (correlation ND) in Japanese Stocks Market. To enhance this signal, we use the Volatility Constrained correlation (VC correlation) method, which led to a significant amplification of this observed tendency. This result has strong implications for increasing predictability of day time return compared to standard correlation. Moreover, the amplified tendency observed for each stock revealed a linear scale relationship between the standard correlation and VC correlation. Taking together, the application of VC methodology to financial trading data overnight enhances the observed correlations which may lead to improve market predictions.

1　はじめに

株式市場や為替市場の価格変動等の分布や相関などに対しては統計的な性質が様々な方面から調べられてる [1-7]。

株式市場には市場が開いている時間（trading period）と市場が閉じている時間（Non-trading period）がある。日本市場が閉まっている時間帯においても、ヨーロッパ市場や米国市場など海外市場は取引をしており、海外のニュースなどが翌日の日本市場に影響を与える。Wang et al[8] は、米国株式市場（New York Stock Exchange (NYSE) stocks）の 1988 年から 2007 年までの 20 年間に対して、の daily return (close-to-close), overnight return (close-to-open) , daytime return (open-to-close) の 3 種類の収益率（return）のそれぞれの組み合わせについての相関などを調べた。この研究により、米国株式市場において overnight return（夜間収益率）と daytime return（日中収益率）の間には弱い負の相関があることが示された。また、Tsai et al.[9] は台湾株式市場（Taiwan Stock Exchange (TWSE)）における daytime return と overnight return の関係を調べた。この研究により、overnight return の振幅が大きい場合、overnight return と次の日の daytime return の正負の符合が逆符合になる確率が高くなることが示された。

我々は、以前の研究 [10] において、金融市場において相関のみならずその影響の方向性まで推定することのできる Volatility Constrained correlation (VC correlation) という方法を導入した。米国市場が日本市場に与える影響が、その逆の方向よりも大きいことがこの方法で定量的に示された。またこの方法は、生物・医学分野の遺伝子データにも応用されている [11]。

2　データ

今回用いたデータは日経 225 インデックスに採用されている 225 社の日足時系列データである。このデータには 2000 年から 2015 年までの日足の初値と終値が含まれる。株式分割があった場合は修正された初値と終値を用いる。

*連絡先：E-mail: ochiai@otsuma.ac.jp
†連絡先：E-mail: nacher@is.sci.toho-u.ac.jp

3 提案方法

日経２２５インデックスの採用銘柄にたいして、daytime return と overnight return の間の相関について調べた。daytime return と overnight return の間には２種類の相関を考えることができる。一つは (Correlation ND (Night to Day)) で、 overnight return と翌日の daytime return の間の相関である。もう一つのタイプは (Correlation DF (Day to Following night) で、daytime return とその後の overnight return の間の相関である（図 1 参照)。

3.1 相関係数

各銘柄 i に対して overnight と daytime の return の間の通常の相関係数を $C^i[N,D]$ とし、daytime と following night の return の間の通常の相関係数を $C^i[D,F]$ とする。

3.2 VC 相関

次に VC 相関 (Volatility Constrained correlation (VC correlation)) を定義する。(詳しい数学的定義は論文 [10] を参照。) 時系列ペアにたいして VC 相関を計算する時、そのうちの 1 つの時系列の標準偏差に着目する。その注目した時系列の return とその平均の差の絶対値が標準偏差よりも大きい時のみの時系列ペアを取り出して相関係数を計算したものが VC 相関である。銘柄ペアには 2 つの時系列があるので、着目することのできる標準偏差は 2 通りある。それゆえ、例えば overnight(N) と daytime(D) の return の 2 つの時系列の間で、２種類の VC 相関係数 $F^i[N,D]$ と $F^i[D,N]$ を計算することができる。ここで、$F^i[N,D]$ では overnight return の標準偏差を用いており、$F^i[D,F]$ では逆に daytime return の標準偏差を用いる。

ここで、この 2 種類の VC 相関の値を比べることにより影響の方向性を推定することができる。例えば $F^i[N,D]$ が $F^i[D,N]$ より大きければ overnight return から daytime return の方向に影響を与えていることがわかる。

3.3 平均相関係数と平均 VC 相関

次に各銘柄 i に対するこれらの相関係数 $C^i[N,D]$ や VC 相関 $F^i[N,D]$ と $F^i[D,N]$ をすべての銘柄に対して平均を取ったものが平均相関係数 $C[N,D]$ や平均 VC 相関 $F[N,D]$ と $F[D,N]$ である。

3.4 相関の相関

また、ある年 y の相関係数 $C^i[N,D](y)$ と次の年 $(y+1)$ の相関係数 $C^i[N,D](y)$ の間の相関係数を $TC[N,D](y)$ とする。ある年 y の VC 相関係数 $F^i[N,D](y)$ と次の年 $(y+1)$ の VC 相関係数 $F^i[N,D](y)$ の間の相関係数を $TF[N,D](y)$ とする。

4 結果

まず、通常の相関係数である $C^i[N,D]$ と $C^i[D,F]$ を調べた。相関係数 $C^i[N,D]$ を計算すると、日本株式市場でほとんどのすべての年において、overnight return と翌日の daytime return が負の弱い相関をしていることが観測された（図 2 参照)。これは、海外（米国、台湾）株式市場の結果と整合している [8, 9]。これに対して、相関係数 $C^i[D,F]$ を計算すると、daytime return と翌日の overnight return の間には明らかな相関は認められないことが分かった。

次に通常の相関係数の代わりに、VC 相関を ND(Night to Day) と DF(Day to Following night) に適用した。ND と DF の両方で VC 相関の絶対値は通常の相関係数より大きくなることが観測された。さらに VC 相関の非対称性より影響の方向性も推定することができることをさきほど説明したが、ND について VC 相関から決まる方向性は、期待通りに夜間セッションから翌日の日中セッションの方向であり、時間の因果律の方向性と一致することが示された（図２（ｂ）参照)。

次に平均的性質ではなくそれぞれの銘柄について通常の相関係数と VC 相関を ND と DF に対して調べた。ND と DF の両方に対して、それぞれの銘柄の VC 相関係数が通常の相関係数の 2 倍ほど増幅されていることがわかった (図 4 参照)。

また、ある年の ND と DF に対する相関係数は、次の年の相関係数と正の相関をすることがわかった（図５参照)。(つまり相関係数の相関が正であるということ。また、VC 相関についても同様の結果になる。) これは各銘柄に対して、ある年の相関の傾向が次の年の相関の傾向に影響与えていることを示している。例えば、ある年のある銘柄の相関 ND が正ならば次の年の相関 ND も正である可能性が高い、

5 終わりに

今回、日経２２５採用銘柄に対して overnight と daytime return (ND)、 daytime return と following over night return (DF) の間の相関について、通常の相関係数と VC 相関を用いて調べた。 まず第一に、overnight と daytime return (correlation ND) の間に弱い負の相

関が観測された。第 2 に VC 相関を用いると、この弱い負の相関が増幅されて観測された。これは、VC 相関を用いることにより翌日の daytime return の予測可能性を高める重要な結果である。また、VC 相関から決まる影響の方向性は時間の因果律と整合性があることが確認された。第 3 にそれぞれの銘柄について VC 相関と通常の相関係数をプロットしてみると傾きが 2 程度の線形の関係があることが分かった。（言い換えると VC 相関で、通常の相関係数が 2 倍に増幅される。）これらの結果から、VC 相関を用いることにより、通常の相関係数よりも予測可能性を改良できることが分かった。

謝辞

T.O. was partially supported by JSPS Grants-in-Aid for Scientific Research (Grant Number 15K01200).

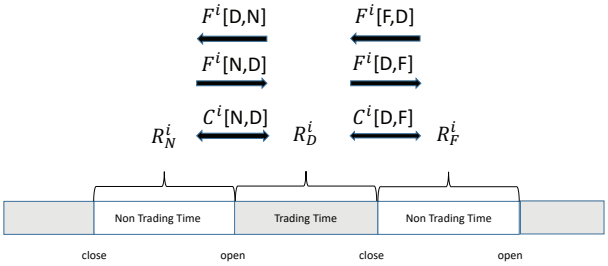


図 1: N(overnight), D(daytime), F(following overnight) のイラスト.

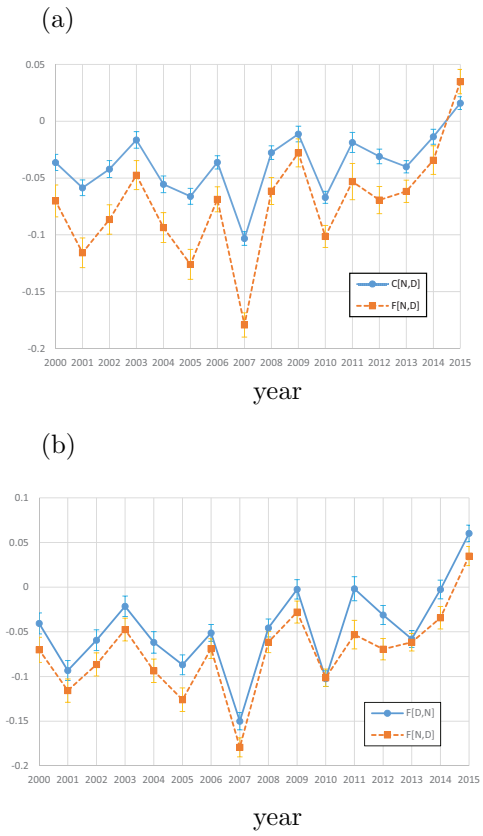


図 2: (a) overnight（N）と翌日の daytime（D）の間の平均相関 $C[N, D]$ と平均 VC 相関 $F[N, D]$ の推移。
(b) overnight と（N）翌日の daytime（D）の間の VC 平均相関 $F[N, D]$ と平均 VC 相関 $F[D, N]$ の推移。両方の図とも error bar は s.e.m である。

参考文献

[1] R. N. Mantenga and H. E. Stanley, *An introduction to Econophysics*, Cambridge University Press, Cambridge, UK (2000).

[2] S. Sinha, A. Chatterjee, A. Chakraborti and B. K. Chakrabarti, Econophysics: An Introduction, Wiley-VCH (2010).

[3] J. M. Pollet and M. Wilson, Journal of Financial Economics, Vol. **96**, Issue 3, 364 (2010).

[4] L. Sandoval and I. D. P. Franca, Physica A **391**, 187 (2012).

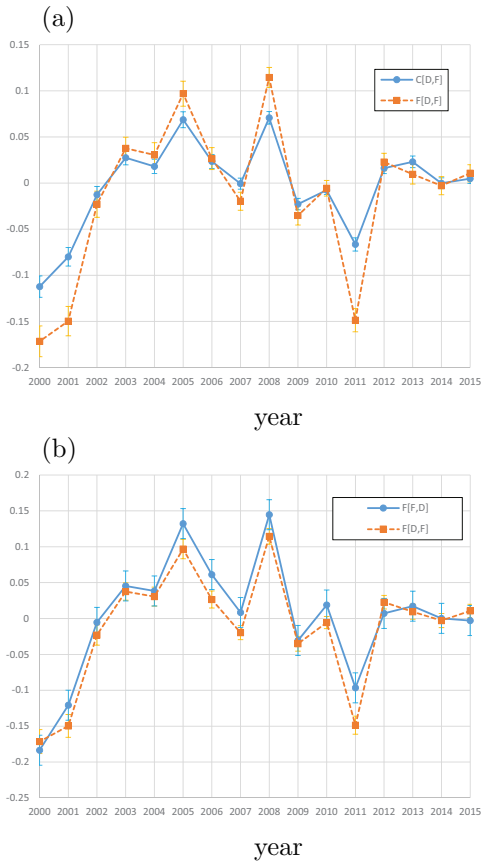


図 3: (a) daytime（D）と following overnight（F）の間の平均相関 $C[D,F]$ と平均 VC 相関 $F[D,F]$ の推移。
(b) daytime（D）と following overnight（F）の間の VC 平均相関 $F[D,F]$ と平均 VC 相関 $F[F,D]$ の推移。両方の図とも error bar は s.e.m である。

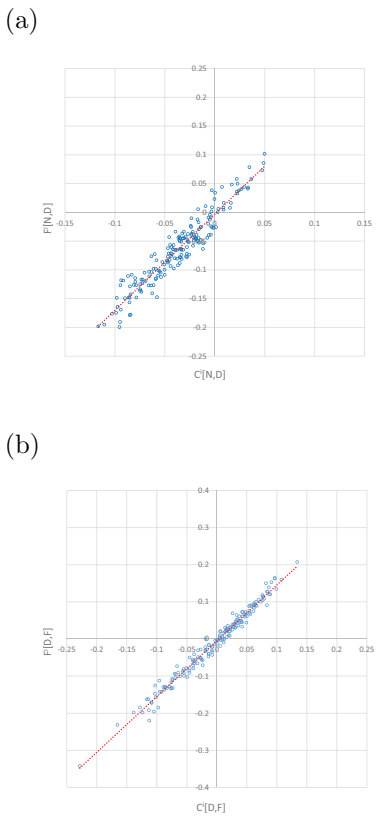
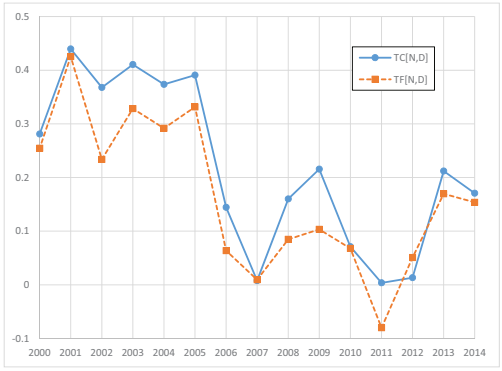


図 4: (a) overnight（N）と翌日の daytime（D）の間のそれぞれの日経２２５採用銘柄 i に対する相関係数 $C^i[N,D]$ と VC 相関係数 $F^i[N,D]$ の散布図。(b) daytime（D）と following overnight（F）の間のそれぞれの日経２２５採用銘柄 i に対する相関係数 $C^i[N,D]$ と VC 相関係数 $F^i[N,D]$ の散布図。両方の図とも相関と VC 相関は２０００年から２０１５年のすべての期間にわたって計算されている。赤点線は regression line である。

(a)



(b)

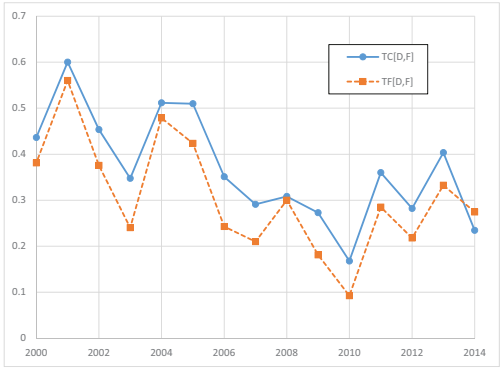


図 5: (a) ある年と次の年の間の相関係数の相関係数 $TC[N, D]$ と VC 相関係数の相関係数 $TF[N, D]$。 (b) 同様に $TC[D, F]$ と $TF[D, F]$。

[5] T. Preis, D. Y. Kenett, H. E. Stanley, D. Helbing and E. Ben-Jacob, Scientific Reports **2**, 752 (2012).

[6] Y.Shapira, D. Y. Kenett, and E. Ben-Jacob, European Physical Journal B. **72**, 657 (2009).

[7] D.-M. Song, M. Tumminello, W.-X. Zhou, R. N. Mantegna, Physical Review E **84** 026108 (2011).

[8] F. Wang, S.J. Shieh, S. Havlin and H. E. Stanley, *Phys Rev E* **79**, 056109 (2009).

[9] K. T. Tsai, J. S. Lih and J. Y. Ko, *Physica A* **391**, 6497 (2012).

[10] T. Ochiai, J.C. Nacher, *Physica A* **393**, 364 (2014) .

[11] T. Ochiai, J.C. Nacher, *Biosystems* **145**, 9 (2016)

B2B 市場における企業ブランドと ROE の関連性

Relationship between corporate brand and ROE in industrial markets

真鍋 友則 [1*]　中川　慧 [2]
Tomonori Manabe[1] Kei Nakagawa[2]

[1] 株式会社 Sansan
[1] Sansan Co,Ltd.
[2] 野村アセットマネジメント株式会社
[2] Nomura Asset Management Co,Ltd.

Abstract:
The importance and necessity of investment in intangible asset value of companies, including ESG investment, has been increasing in recent years. Corporate brands are an important element of intangible assets. However, especially in B2B companies, there are many unclear points regarding the relationship between corporate brand value and business performance. In this study, we quantitatively evaluated the relationship between B2B corporate brands and performance using a new large-scale corporate brand survey index created from a business card exchange network. We examined the relationship between the corporate brand index and ROE for B2B companies. As a result, we found that companies with a high corporate brand index had a significantly higher profit margin, a significantly lower turnover rate and no significant association with financial leverage. These results show that even in B2B companies, companies with high corporate brands have implemented a differentiation strategy that allows a high margin range, while companies with low brands have adopted cost leadership strategies with high turnover rate.

1　はじめに

1.1　無形資産投資と企業ブランド

企業価値は企業の有形資産の価値と無形資産の価値の和によって決定される. 無形資産は知的財産, 商標, 技術や知識の集積, 人的資源, 組織風土, 企業や商品に関する風評など, 財務指標には表れにくい, 企業の有する無形の資源である [9]. 企業価値を創出する資産の面からみると, 20 世紀は土地, 建物, 工場, 機械などの物理的な資産や現金や有価証券などの金融資産といった有形資産の時代であった. しかし, 21 世紀は, 目に見えないような無形資産が企業価値の創出の主な構成要素になってきていると言われている [25]. 近年では, 経済産業省の「持続的成長に向けた長期投資 (ESG・無形資産投資) 研究会」のレポート ”伊藤レポート 2.0”が, 英国と米国においてこのような無形資産への投資総額が有形資産への投資額を上回っていること, また, 国連による「責任投資原則」を契機とし, 長期投資に ESG (環境、社会、ガバナンス) 等の「非財務情報」を重視する動きが拡大していることを指摘している [24].

ブランドはこのような無形資産の要素である [11, 24]. ブランドの価値は, D. Aaker によれば, ”商品やサービスを通して消費者や企業へ提供される価値に付加される (または差し引かれる), ブランド, その名称やシンボルに結びついた一連の資産 (または負債)” として定義されている [1]. 企業ブランドは, 組織全体に対する想起から形成され, そこには, 企業の資産や能力, 人材, 価値感や使命, 行動基準, 社会参画, 業績などが含有されている [2]. 強い企業ブランドは商品や企業の競争優位を創出し, 高い利益率, 新製品への拡張, 顧客ロイヤリティの向上, 従業員満足度の向上, 差別化の継続的実現など, 種々の利益を企業にもたらすとされている [6]. さらに, 資産として定量化されたブランド価値は, 株式市場における評価, 例えば時価総額 [5] や, 株価のリターン [16] とも関連していることが明らかにされており, 投資判断においても重要なファクターとされる.

*連絡先： Sansan 株式会社
107-0061 東京都港区北青山 3 - 5 - 29 One 表参道 6F
E-mail: tom.manabe@gmail.com

1.2 B2B 企業ブランド

企業ブランドは歴史的に, 主として商品に上乗せされる付加価値として位置付けられてきた. その一方で, 購買意思決定プロセスが合理的であるとされる B2B ビジネスにおいては, 感情的判断と結びつくようなブランディング戦略は, それほど重要な要因ではないと考えられてきた [14]. しかし, B2B 企業ブランドが, 購買意思決定へ影響するとともに, ステークホルダーとの間の感情的な結びつきや信頼の醸成という点で重要であることが, 研究によって明らかにされつつあり [10, 18], B2B 企業ブランディングに関する実証的研究の数は年々増加傾向にある [20]. 一方で, B2B 企業ブランドの定量化や, 利益への影響について, まだ明らかにされていない点も多く残されている [14].

B2B と B2C ではブランドの構成要素が異なるとされる. Keller が提唱した CBBE ピラミッド [12] を B2B の文脈で捉え直し, 再構成したモデルが Kuhn らによって提案された [13]. Kuhn らによれば, B2B 企業においては個々の製品よりも, 企業全体のブランドがより重要となる. また, イメージの次元は B2B ブランドでは除外され, 企業に対する風評 (reputation) が, ブランド構成要素として導入されている. さらに, セールスパーソンやサービス担当者などとの関係性の次元も重要なファクターとして, ブランド形成の上位に導入されている [13, 14]. このような, 企業全体への認知や風評が B2B ブランドで重要視される背景として, 次のような購買プロセス上の特徴が指摘されている. すなわち, B2B の購買プロセスは一般に, 様々な立場の意思決定者が複数人存在しているプロセスである. そのようなマルチ・ディシジョン・プロセスでは最大公約数的・リスク回避的な合意形成がなされやすく, 風評の良い企業, 高い認知の企業は有利になる [23].

これらの知見は, 企業全体の風評や, セールスパーソンやアフターセール担当者などについての風評が, 製品やサービスのパフォーマンス・価格・納期などの合理的な判断軸に加え, 購買決定要因となり得ること, とりわけ, 購買リスク低減要因として機能することを示唆している. ただし, その知見は一部の業界のインタビューや, B2B 企業の購買プロセスの特徴の考察に基づいて得られたものであり, 大規模な定量調査によって一般的に実証された結果ではない.

B2B 企業に関するブランド印象や風評は, B2C 企業と比べて次のような理由から定量調査が困難である. ある B2B 企業について, その業態やサービスに関して一定以上の知識を有する者は実際に商取引のある人々または近しい業種の関係者に限定される. そのため B2B 企業ブランドと企業の業績に関する実証的研究は, これまで十分に行われて来なかった.

本稿では, この問題にアプローチするため, 名刺アプリ Eight のネットワークデータを応用した企業ブランド印象調査指標である BBES (B2B Engagement Score, Sansan 株式会社) [19] を用い, これを B2B 企業ブランド指標として, 企業業績 (ROE) との関係を分析する.

BBES は, 調査対象企業の名刺を保有しているユーザーのみを調査母集団とし, 回答者の評点により算出されたスコアである. 名刺を保有していることから, 企業内部の人間とビジネス上の接点を有していることが担保されており, したがってその調査対象集団による企業ブランド評価指標は, ビジネスとの関連におけるブランド印象を反映していると期待される. 本スコアは, 大規模な名刺交換ネットワークの活用により, 1600 社余りの企業についての評価を, 規模・業種横断的に調査しているという利点も有しており, このスコアを活用することにより, 特定の企業や業種に偏らないマクロレベルでの実証的分析が可能になると期待できる.

1.3 ブランドと企業業績

企業ブランド向上のために, どれだけ投資すべきか, あるいはどのような戦略を取るべきかなどの意思決定のために, ブランドと企業業績の関連性についての実証的分析の知見が重要となる. ブランドと企業業績との関係性の事例として, ブランド価値と ROA の関係についての研究が報告されている.

例えば, 風評の良い企業は, 高い ROA をより長く持続させることができることが示されている [18]. また, ブランド価値が, 企業の将来の ROA に正の影響を及ぼすことも報告されている [7, 15]. さらに, ブランド価値における企業の「社会的責任」が, ROA と正の関連性をもつことが示されており, 企業の CSR 評価が ROA と関連することも明らかにされている [22].

また, ブランド価値と利益率の間に関連があることも報告されている. その報告によれば, アメリカの自動車市場の調査において, ブランド価値要素の「差別化」と「知識」が, マージンの割合と正の関連性を有している [21].

上記の報告はいずれも, B2B 企業に限定した分析ではない. B2B 市場におけるブランド価値と企業業績に関しては, Homburg らの研究により, ブランド認知と売上成長の間に正の関連性があることが明らかになっているが, ただし, 企業の売上成長性はアンケートデータに基づいており, 実際の財務指標は用いられていないなど, 定性的な関連性の指摘に留まっている [10].

2 本稿の位置付けとリサーチデザイン

ブランドが企業業績に影響することは上記のような研究で示されており, その機能は徐々に明らかにされてきているが, B2B 企業に焦点を当てた研究は限定的である. そこで本稿では, B2B 企業ブランドが有している役割を, 企業業績 (ROE) と関連付け, 以下の仮説に基づいて検証する.

B2B 企業間取引では, 様々な立場の意思決定者が複数人存在しているプロセスであるため取引に伴うリスクが回避される傾向にあるとされる [23]. また, B2B においてブランドが高い企業は, 取引に関するリスクが小さい企業であるという報告もある [4]. これらのことから, ブランド価値が高い B2B 企業は, リスクを価格に反映させる必要性が低く, 相対的に高価格での取引成立を可能し, そのことにより高い利益率を実現させると予想することができる. 一方, ブランド価値が低い企業は, 取引リスクがもたらす低い利益率を補う形で, 回転率を高めるコストリーダーシップ戦略を取っていると予想される. また, ブランドが高い企業は, 負債調達がより容易になり, 財務レバレッジが高いと予想される. 以上の予想より, 次の仮説を立てる (図 1).

H1 ブランド指標の高い B2B 企業は, 利益率が高い.

H2 ブランド指標の低い B2B 企業は, 回転率が高い.

H3 ブランド指標の高い B2B 企業は, 財務レバレッジが高い.

本研究は, 上記仮説を検証し, B2B 企業のブランドが有している役割を企業業績と関連付けて明らかにすることを目的とする.

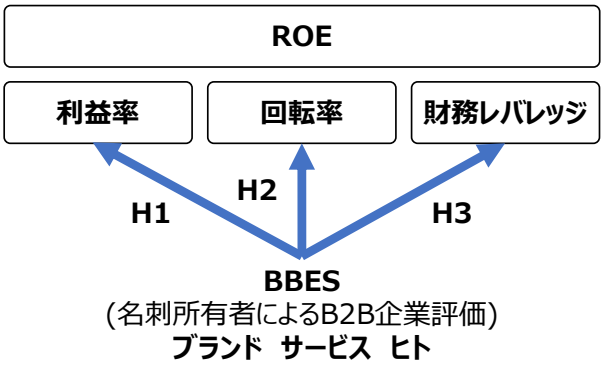


図 1: リサーチデザイン:BBES スコアは ROE を構成する 3 つの要因とどのような関係にあるのか

2.1 ROE 分解

ROE の企業間での比較分析の代表的方法として, いわゆるデュポン分解がファイナンスの実務においても広く利用されている. ある企業の ROE = 利益/純資産 ,財務レバレッジ = 総資産/純資産, ROA = 利益/総資産, 利益率 = 利益/売上, 回転率 = 売上/総資産 と定義する. すると, ROE と各変数間の関係は次式のように分解した形で記述される.

$$\begin{aligned} ROE &= \text{財務レバレッジ} \times ROA \\ &= \text{財務レバレッジ} \times \text{利益率} \times \text{回転率} \end{aligned}$$

デュポン分解を考える利点として, ROE が大きい (小さい) 要因が, 財務レバレッジ, 利益率, 回転率という 3 つのどの項にあるのかを特定し, より深い考察が可能になることが挙げられる.

財務レバレッジは資金調達を負債で行うか, 株主資本で行うかのバランスを表している. ROE は分母が株主資本であるため, 負債が多いと分母が小さくなるため, ROE は上昇する. 一方で, 負債での調達を増やすと一般に信用リスクが上昇し, デフォルト等の悪影響が生じる.

利益率と回転率は調達した資金の運用による効率を表す. 利益率は, 売上に占める利益の割合であり, 付加価値の大きさを表し, 回転率は, 資産に対してどれだけの売上をあげたかという効率性を表す. 利益率と回転率は一般にトレードオフの関係にあるとされる. Porter によると高付加価値な差別化戦略をとると利益率は高くなるが, 回転率は低くなる [17]. 一方で, 薄利多売というコストリーダーシップ戦略を取ると回転率は高くなるが, 利益率は低くなる. なお, 本研究では利益率として企業の本業の利益率を計測するという観点から, 売上高営業利益率を用いる.

2.2 BBES (B2B Engagement Score)

B2B 企業ブランドの指標として分析に用いる BBES の説明を行う. Sansan 社の BBES は, 半年に一回, 調査対象企業の名刺を有する 名刺アプリ Eight の ユーザーにランダムに調査票をメール送付し, 任意のアンケート結果を回収することで作成される. 調査母集団を名刺所有者に限定することにより, 日本の全労働人口からのランダムサンプリングと比べ, その企業の人と実際に面識があり関連度のより高い人たちを対象とした調査が実現できるという特徴がある. 一方で, 例えばインターブランド社のようなブランド評価とは異なり, アンケートによる評価は金額ベースのものではなく, 0 から 10 までの離散的な評価である. これに関し, 鳥居 [27] によるとブランドを金額として評価したり, 財務諸

表の枠内で検討したりする必要性は必ずしもない．ブランドはイメージや抽象的なもので捉えても問題ないが, 価値関連性を考える際にブランドを比較するため少なくとも数値的なデータである必要はある．

調査はこれまで 2018 年の 5 月と 11 月に行われており, 本稿では 11 月調査を分析に利用した．対象企業数と, そのうち主な業態が B2B 事業である企業数[1], 調査人数を表 1 に記す．

表 1: BBES 2018 年 11 月調査企業数と回答者数

調査対象企業数	1,388
上場企業数	582
上場企業のうち B2B 企業数	323
1 企業あたりの対象調査人数	1,500
平均有効回答者数 (最小値-最大値)	154 (87-218)

アンケートは「ブランド」, 「サービス」, 「ヒト」の 3 項目に関する意識調査から成り, 回答者は,

- ブランド: x 社のブランドイメージは魅力的だと思いますか?
- サービス: x 社の製品・サービスは自社/社会に有用だと思いますか？
- ヒト: x 社の人は好印象だと思いますか？

の 3 つの質問に対し, 0 - 10 の 11 段階の評点をつける．1 人の回答者が回答可能な企業数は, 最大 3 企業に限定されている．

ある企業に対する回答者の得点平均値を, その企業の BBES スコアとし, 以降の分析ではこの数値を用いる．BBES スコアの平均値 (標準偏差) は表 2 のようであった．参考として, B2B 企業に関して, スコアの高い企業 Top 10 を表 3 に記す．なお, 以降の分析は, B2B と分類される企業に限定して行なっている．

表 2: 2018 年 11 月調査 BBES スコアの分布

	平均値	標準偏差
ブランド	6.49	0.75
サービス	6.56	0.67
ヒト	6.50	0.50

2.3 BBES と財務データ

2018 年の企業の各種財務指標およびファイナンスデータは野村総合研究所の分析データ・サービスにより取得した．BBES と財務指標間の相関係数表を表 4 に示す．BBES スコア の各項目間の相関は相互に正に高い ($p < 0.01$) ．3 項目の中ではヒトスコアがやや, 他の 2 項目との関連が弱い．BBES スコア と企業の売上高には有意な正の相関があり ($p < 0.01$), 規模の大きい企業の方がスコアが高いことがわかる．ROA や PBR と BBES の間には負の相関関係がみられ, 特にヒトスコアとの間に有意性がみられる ($p < 0.01$)．ただし, PBR と ROA は売上高との間にも負の相関関係があるため ($p < 0.05$), 売上高が共変量として働いているとも考えられる．

BBES スコアを対数変換した売上高と, 業種[2]で回帰したモデルを式 1 に記す．

$$BBES_{ik} = \beta_{1k} ln(Sales_i) + \sum_{j=1}^{10} \beta_{2kj} I_{ji} + \varepsilon_{ik} \quad (1)$$

$$relativeBBES_{ik} = \varepsilon_{ik} \quad (2)$$

ここで, $BBES_{ik}$ は企業 i の項目 k (ブランド, サービス, ヒト) のスコアを表す．$Sales_i$ は企業 i の売上高, I_{ji} は企業 i が業種 j に属しているかどうかのダミー変数で, 属している場合は 1, そうでない場合は 0 をとる．BBES スコア k に対する回帰式 (1) の残差 ε_{ik} は, 売上高と業種から推定される BBES スコアの期待値からの乖離を表している．この ε_{ik} を, 規模・業種調整済み BBES として, relative BBES と名付ける (式 2)．relative BBES の Top 10 ランキングは表 5 のようになる．高順位が大企業で占められていた BBES (表 3) と比べ, relative BBES では規模の要素を排したブランド力をもつ企業が抽出されているといえる．また, relative BBES のブランド, サービス, ヒトスコア と ROE, ROA, および PBR との相関は表 4 のようになる．relative BBES スコア同士は BBES スコアと同様に相互に相関があり, 規模や業種とは独立な要因間においても, 互いに相関していることが分かる．ROE, ROA との相関値については, 統計的有意を示す組み合わせはなく, 売上高を統制することにより, ヒトスコアと ROE の負の相関値は 0 レベルを示している．relative BBES のブランドと PBR の間には弱い正の相関が見られる ($p < 0.1$)．このことは, 株式市場における企業評価と規模や業種の影響を除した周囲のブランド評価である relative BBES の間に関連性があることを示す．

3 実証分析

ここでは, 仮説 H1, H2, H3 を検証するための実証分析を行う．2.3 節で示したように, BBES と売上高で示

[1] 企業の業態分類は Baseconncet 株式会社の提供する情報に基づく

[2] 業種分類は, 証券コード協議会の定める業種分類のうち, 大分類 (10 分類) を用いた [26].

表 3: B2B 企業の 2018 年 11 月調査 BBES ランキング

	BBES		
順位	ブランド	サービス	ヒト
1	株式会社デンソー	株式会社デンソー	さくらインターネット株式会社
2	鹿島建設株式会社	日本製鉄株式会社	鹿島建設株式会社
3	三井物産株式会社	ＳＭＣ株式会社	株式会社大林組
4	株式会社小松製作所	株式会社クボタ	ＡＧＣ株式会社
5	ＤＭＧ森精機株式会社	株式会社安川電機	株式会社豊田自動織機
6	株式会社大林組	株式会社村田製作所	川崎重工業株式会社
7	株式会社村田製作所	ＡＧＣ株式会社	株式会社クボタ
8	株式会社クボタ	住友化学株式会社	コニカミノルタ株式会社
9	清水建設株式会社	株式会社ゼンリン	理想科学工業株式会社
10	住友化学株式会社	東レ株式会社	戸田建設株式会社

表 4: BBES と売上高, ROE, ROA の相関係数

		BBES						
		ブランド	サービス	ヒト	売上高	ROE	ROA	PBR
BBES	ブランド	-	-	-	-	-	-	-
	サービス	0.91***	-	-	-	-	-	-
	ヒト	0.85***	0.84***	-	-	-	-	-
	売上高	0.48***	0.37***	0.34***	-	-	-	-
	ROE	0.00	−0.05	−0.07	0.01	-	-	-
	ROA	−0.07	−0.09	−0.19***	−0.11*	0.82***	-	-
	PBR	−0.11*	−0.14**	−0.22***	−0.13**	0.33***	0.44***	-
Note:					*p<0.1; **p<0.05; ***p<0.01			

表 5: B2B 企業の 2018 年 11 月調査 relative BBES ランキング

	relative BBES		
順位	ブランド	サービス	ヒト
1	サイボウズ株式会社	株式会社ゼンリン	さくらインターネット株式会社
2	さくらインターネット株式会社	アズワン株式会社	サイボウズ株式会社
3	トラスコ中山株式会社	サイボウズ株式会社	理想科学工業株式会社
4	アズワン株式会社	トラスコ中山株式会社	株式会社アマナ
5	三井物産株式会社	さくらインターネット株式会社	山九株式会社
6	株式会社ゼンリン	株式会社ミロク情報サービス	アズワン株式会社
7	株式会社アマナ	因幡電機産業株式会社	株式会社船場
8	鹿島建設株式会社	日本工営株式会社	株式会社乃村工藝社
9	株式会社ラック	株式会社アマナ	株式会社たけびし
10	株式会社乃村工藝社	ＳＭＣ株式会社	鹿島建設株式会社

される企業規模には相関があり, 交絡要因として働く恐れがある. このため, 売上高と業種に相関する成分を取り除いた relative BBES と, ROE コンポーネントの間の相関を調べる.

relative BBES と ROE コンポーネントの要約統計量を表 7 に記した. また, relative BBES と ROE コンポーネント間の相関係数および t 検定の結果を表 8 に示している.

利益率に対して, relative BBES ブランドとサービスは, 有意に正の相関係数値を示しており, これは仮説 H1 を支持する結果となっている. すなわち, ブランド評価が高い B2B 企業はマージン率も高い傾向にあると言うことができる. 一方で, 回転率に対しては, 統計的有意に負の相関係数値を示しており, これは H2 を支持する結果であることがわかる. すなわち, ブランドやサービス評価の相対的に低いような B2B 企業では,

表 6: relative BBES と売上高, ROE, ROA の相関係数

		relative BBES					
		ブランド	サービス	ヒト	ROE	ROA	PBR
relative BBES	ブランド	-	-	-	-	-	-
	サービス	0.84***	-	-	-	-	-
	ヒト	0.70***	0.69***	-	-	-	-
	ROE	0.09	0.04	−0.03	-	-	-
	ROA	0.09	0.06	−0.09	0.78***	-	-
	PBR	0.17*	0.13	0.02	0.33***	0.44***	-

Note: *p<0.1; **p<0.05; ***p<0.01

表 7: ROE コンポーネントと relative BBES の要約統計量

	平均	標準偏差	中央値	最小値	最大値	歪度	尖度
利益率	0.08	0.07	0.06	−0.12	0.48	2.18	9.12
回転率	1.13	0.56	0.97	0.16	3.21	1.47	2.54
レバレッジ	2.29	0.96	1.98	1.05	7.30	1.56	3.57
relative BBES ブランド	0.00	0.47	0.00	−1.60	1.73	−0.01	0.55
relative BBES サービス	0.00	0.42	0.36	−1.38	1.47	−0.13	1.03
relative BBES ヒト	−0.00	0.33	0.03	−0.90	1.43	−0.15	1.00

表 8: relative BBES と ROE コンポーネントの相関係数

relative BBES	利益率	回転率	レバレッジ
ブランド	0.22***	−0.19***	−0.05
サービス	0.19***	−0.19***	−0.04
ヒト	−0.03	−0.09	0.03

Note: *p<0.1; **p<0.05; ***p<0.01, 企業数=272

コストリーダーシップ戦略が取られている傾向にある. relative BBES 項目の内, ヒトについては, このような有意な関係性が見られない. これは, 企業の社員に対する評判と, 企業の利益率や効率性との間に, 特に関連が見られないことを示している.

レバレッジに対してはどの項目も有意な相関関係は見られず, H3 を支持する結果は得られなかった.

4 考察

本研究は, ROE のデュポン分解から見られる B2B 企業のブランド戦略と, B2B 企業周囲からのブランド印象の間に有意な関連があることを, 大規模調査データで示したものである.

ブランド戦略とコストリーダーシップ戦略は戦略の違いであり, 企業の将来の収益性に対し, どちらが有利であるとは言えるものではないとされる [8]. したがってこの分析からのみでは, 必ずしもブランドの向上が企業のリターンに結び付くとは言えない. ただし, B2B 企業が, 利益率の向上を目指す場合, 「ブランド力」向上の施策や投資, マネージメントの正当性を示唆する結果であるといえる.

B2C においては, 製品のブランド価値と, 収益プレミアムの関係が示されてきた [3]. 今回の結果は, 同様な傾向が, 従来はブランディングがあまり寄与しないと言われてきた B2B 企業においても見られることを示したものである. ただし, それを可能にしているメカニズムは, Kuhn ら や Webster ら の B2B ブランド形成に関する考察を参照した場合, B2C とは異なっていると推察される [13, 23]. B2B においては, ブランドは B2C と異なり消費者の感情的な結びつきやイメージ上の親和性によってもたらされるものではなく, 複数の意思決定の中で, 信頼感や低リスクと言った要因で形成される. したがって, B2B 企業のブランド印象が利益率と関連する背景としては, 次のようなメカニズムの想定がより妥当と思われる. すなわち, ブランドの印象評価が高く, 風評の良い B2B 企業は, 取引に伴うリスクが低下するため, 高いマージンが許容される.

本分析の限界として, 同時点の関連性の分析に止まっており, 高いブランド印象が高い利益率を可能にしているかどうかを分析的に示すためには, より詳細な時系列

分析が必要になる. 今後調査データが蓄積されることで, そのような分析も可能になると考えられる.

次の研究課題として, B2B 企業ブランドの決定要因は何か, という問題がある. BBES には企業についてのフリーコメントも付与されているため, そのコメントを分析することで, この問題にアプローチできると考えている.

また, relative BBES と利益率・回転率の回帰においては, その相関係数は有意であるが残差も大きい. これは, 周囲からのブランド評価は低いが利益率が高い企業や, その逆の傾向の企業の存在を示しており, 利益率をブランド要因とそれ以外の要因に区分できる可能性があることを示唆している. 今後, この区分が, 将来の利益率の変化や持続と関係があるかどうか, すなわち, ブランドに支えられた利益率とそうではない利益率が, それぞれ時間的にどのように変化していくのかを分析したい. そのような分析を通し, B2B 企業ブランドと企業の持続可能性の関係について, 今後明らかにしていく.

参考文献

[1] David A. Aaker. *Managing Brand Equity*. Business Economics, 1991.

[2] David A. Aaker. Leveraging the corporate brand. *California Management Review*, No. 46, p. 3, 2004.

[3] Kusum L. Ailawadi, Donald R. Lehmann, and Scott A. Neslin. Revenue premium as an outcome measure of brand equity. *Journal of Marketing*, No. 67, pp. 371–385, 2003.

[4] Klaus Backhausa, Michael Steinerb, and Kai Lggera. To invest, or not to invest, in brands? drivers of brand relevance in b2b markets. *Industrial Marketing Management*, No. 40, pp. 1082–1092, 2011.

[5] Mary E. Barth, Michael B. Clement, George Foster, and Ron Kasznik. Brand values and capital market valuation. *Review of Accounting Studies*, Vol. 3, p. 4168, 1998.

[6] Scott Davis. Brand asset management: how businesses can profit from the power of brand. *Journal of Consumer Marketing*, No. 19, p. 4, 2002.

[7] Li Li Eng and Hean Tat Keh. The effects of advertising and brand value on future operating and market performance. *Journal of Advertising*, No. 36, pp. 91–100, 2007.

[8] Patrica M. Fairfield and Terilombardi Yohn. Using asset turnorver and profit margin to forecast changes in profitability. *Review of Accounting Studies*, No. 6, pp. 371–385, 2001.

[9] Richard Hall. The strategic analysis of intangible resources. *Strategic Management Journal*, No. 13, pp. 135–144, 1992.

[10] Christian Homburg, Martin Klarmann, and Jens Schmitt. Brand awareness in business markets: When is it related to firm performance. *Intern. J. of Research in Marketing*, No. 27, pp. 201–212, 2010.

[11] Hiroyuki Itami and Thomas W. Roehl. *Mobilizing Invisible Assets*. Harvard University Press, 1991.

[12] Kevin Lane Keller. *Strategic brand management: building, measuring, and managing brand equity*. Pearson, 4 edition, 2012.

[13] Kerri Louise Kuhn, Frank Alpert, and Nigel K. Ll. Pope. An application of keller ' s brand equity model in a b2b context. *Qualitative Market Research*, No. 11, pp. 40–58, 2008.

[14] Sheena Leek and George Christodoulides. A literature review and future agenda for b2b branding: Challenges of branding in a b2b context. *Industrial Marketing Management*, No. 40, pp. 830–837, 2011.

[15] Natalie Mizik. Assessing the total financial performance impact of brand equity with limited time-series data. *Journal of Marketing Research*, No. 51, pp. 691–706, 2014.

[16] Natalie Mizik and Robert Jacobson. The financial value impact of perceptual brand attributes. *Journal of Marketing Research*, No. 45, pp. 15–32, 2008.

[17] Michael E Porter. *Competitive strategy: Techniques for analyzing industries and competitors*. Simon and Schuster, 2008.

[18] Peter W. Roberts and Grahame R. Dowling. Corporate reputation and sustained superior financial performance. *Journal of Business Research*, No. 69, pp. 2664–2677, 2002.

[19] Sansan Inc. BBES ホームページ. https://bbes.sansan.com/. Accessed: 2019-01-15.

[20] Zahra Seyedghorban, Margaret Jekanyika Matanda, and Peter LaPlaca. Advancing theory and knowledge in the business-to-business branding literature. *Journal of Business Research*, No. 69, pp. 2664–2677, 2016.

[21] Florian Stahl, Mark Heitmann, Donald R. Lehmann, and Scott A. Neslin. The impact of brand equity on customer acquisition, retention, and profit margin. *Journal of Marketing*, No. 76, pp. 44–63, 2012.

[22] David Han-Min Wang, Pei-Hua Chen, Tiffany Hui-Kuang Yu, and Chih-Yi Hsiao. The effects of corporate social responsibility on brand equity and firm performance. *Journal of Business Research*, No. 68, pp. 2232–2236, 2015.

[23] Frederick E. Webster Jr and Kevin Lane Keller. A roadmap for branding in industrial markets. *Journal of Brand Management*, No. 11, pp. 388–402, 2004.

[24] 経済産業省. 伊藤レポート 2.0　持続的成長に向けた長期投資　(esg・無形資産投資) 研究会報告書. Technical report, 2017.

[25] 寺本義也, 原田保. 無形資産価値経営: コンテクスト・イノベーションの原理と実践. 生産性出版, 2006.

[26] 証券コード協議会. 業種. https://www.jpx.co.jp/sicc/sectors/01.html. Accessed: 2019-04-10.

[27] 鳥居宏史. ブランド価値測定に関する一考察. 明治学院論叢, No. 688, pp. 67–82, 2003.

GCNによる取引関係グラフからの企業の特徴量抽出

Feature extraction of companies from business relationship graphs by GCN

森 正和[1]　與五澤 守[1]　工藤 剛[2]

Masakazu Mori[1], Mamoru Yogosawa[1] and Tsuyoshi Kudo[2]

[1]株式会社日本総合研究所
[1]The Japan Research Institute, Limited
[2]株式会社三井住友銀行
[2] Sumitomo Mitsui Banking Corporation

Abstract: Financial institutions have a large amount of data on money transfer among companies. With these data, they can construct graphs that represent the business relationships of the companies. If they can predict a rating of a company's repayment capacity from the graphs, they can make more appropriate loan decisions and find new customers. In this study, we propose a method applying extended Graph Convolutional Networks to business relationship graphs representing the continuity of transactions to predict the ratings. As a result of applying this method to actual data, it was indicated that this method automatically extracted features necessary for the prediction of the rating.

１．はじめに

金融機関は企業に融資を行う際に，主にその企業の財務諸表（貸借対照表・損益計算書等）を参考にしてその企業の返済能力を算定し，融資実行の判断や融資額・利率等の決定を行う．しかし，中小企業や新興国の企業においては正確で信頼性の高い財務諸表を入手できない場合があり，結果的に適切な融資判断が行えない場合がある．また，金融機関がまだ融資したことがない企業の中から融資先として有望な企業を絞り込むために返済能力が低い企業を除外するには，事前に全ての企業の財務諸表が必要となるが，株式を公開していない企業の財務諸表は一般的に入手困難であり，従って返済能力が低い企業を除外することも困難である．

一方で，金融機関は企業の預金口座を保有しているため，企業間で取引がなされた際に発生する送金の情報（以下，この情報を「送金明細」と呼ぶ）を保有している．この送金明細から企業の返済能力を予測することができれば，金融機関が融資判断を行う際に企業から入手した財務諸表の信頼性を判定することが可能となる．また，金融機関が融資先として有望な企業を検索することも可能となる．

これまでデフォルト（債務不履行）の予測に関しては[１]や[２]をはじめ，多くの先行研究がなされてきたが，それらの先行研究はいずれも財務諸表を元にして作成した財務指標を説明変数としてデフォルトを予測するものである．本研究は財務諸表を使用しない点がこれらの先行研究と異なる．

また，財務諸表を使用しない先行研究として，[３]や[４]では入出金明細を元に数千の特徴量を設計・加工し，勾配ブースティング決定木等の機械学習モデルを用いてデフォルトの予測を行う手法が提案されている．本研究はこのような多くの特徴量を設計・加工する必要がない点が先行研究と異なる．

本研究では財務諸表や入出金明細の代わりに送金明細を使用する．送金明細は企業の取引関係を表しており，ここから企業をノード，企業間の取引関係をエッジとしたグラフを構築することが可能となる．近年，深層学習をグラフに応用した Graph Convolutional Network，及び，その派生手法が様々なグラフ分析タスクにおいて高い成績を収めている．（[５], [６], [７]）本研究ではこの Graph Convolutional Network を適用し，財務諸表に頼ることなく，企業の取引関係を表すグラフから自動的に特徴量を抽出し，企業の格付を予測する手法を提案する．

２．使用データと提案手法

２．１．使用データ

本研究で使用するデータは特定の一金融機関が保有する送金明細データと企業の格付データである．

送金明細データは振込・振替等の為替情報を元に

作成された企業間の送金情報データである．送金明細データは送金企業・受取企業・日付・金額の４つの情報が含まれる．企業が同じ金融機関に複数の預金口座を持っている場合や，他行にも預金口座を持っている場合には，企業名等の情報を元に同一企業として名寄せを行う．他行の預金口座間で行われた送金情報は送金明細データに含まれない．本研究では金額が 10 万円以上の送金明細のみを使用する．

企業の格付データは企業の返済能力を表す順序尺度を持ったカテゴリー変数である．格付データは各企業の決算期に作成された財務諸表等の情報を元に金融機関によって設定される．各企業の格付データは少なくとも 1 年に一度の頻度で見直されるが，債務不履行等特定のイベントがあった場合は随時見直される．格付が設定された企業は送金明細データに含まれる企業のごく一部である．

本研究では 3 年間の送金明細データを元に，その１年後に設定されていた格付が特定カテゴリー以下となっている返済能力が低い企業を予測する．なお，今回使用したデータにおいて，格付が特定カテゴリー以下の企業の割合は，訓練データは約 19%，テストデータは約 17%であった．

表１は本研究で使用した訓練データとテストデータの期間を示している．

表１：訓練データ・テストデータの期間

	送金明細	格付
訓練データ	2013 年 4 月～2016 年 3 月	2017 年 3 月時点の格付
テストデータ	2014 年 4 月～2017 年 3 月	2018 年 3 月時点の格付

２．２．取引関係グラフ

次に，送金明細データから取引関係グラフを作成する手順を説明する．

まず送金明細データを半期ごとに送金企業・受取企業を単位にして集計する．例えば，2013 年度上期に A 企業から B 企業に 100 万円が送金された，といった形で送金明細を集計する．

この半期ごとに集計されたデータを元に，格付の予測を行う企業（以下，「予測対象企業」と呼ぶ）ごとに，以下に示す手順で，販売先企業や仕入先企業（以下，両者をあわせて「取引先企業」と呼ぶ）との取引関係を示す有向グラフを構築する．

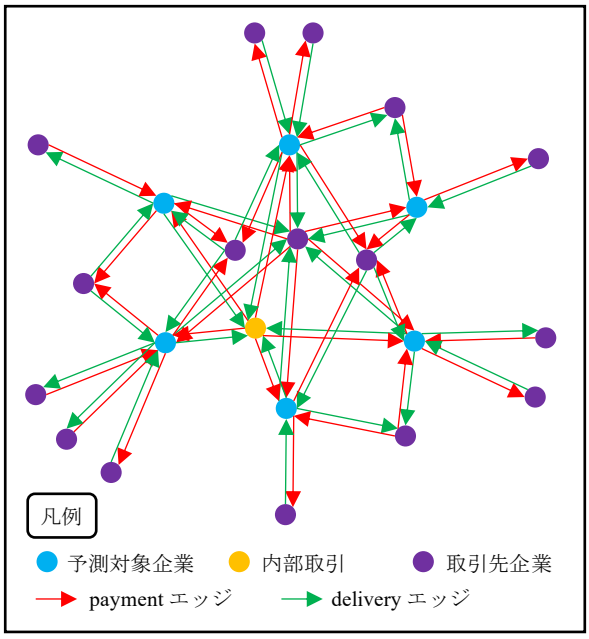
まず予測対象企業の３年間における各半期を示す６つのノードを作成する．次に，予測対象企業の内部取引を示すノードと取引先企業を示すノードを作成する．内部取引と取引先企業のノードは半期ごとに分割せず，３年間で共通の１つのノードとする．そして，予測対象企業の各半期に内部取引や取引先企業と取引が行われていた場合は，送金企業・受取企業間に双方向のエッジを結ぶ．送金方向のエッジを payment エッジ，逆方向のエッジを delivery エッジと呼ぶ．これらのエッジとは別に各ノードにはセルフループエッジを追加する．

ノード特徴量にはノードの種類（予測対象企業の各年度半期，及び，内部取引，取引先）を示す８次元のワンホットベクトル $\boldsymbol{h}_i^{(0)} \in \mathbb{R}^8\ (i = 1, \cdots, N)$ を設定する．ここで N は取引関係グラフのノード数を示す．

payment と delivery のエッジ特徴量 $p_{i,j} \in \mathbb{R}, d_{i,j} \in \mathbb{R}\ (i = 1, \cdots, N, j = 1, \cdots, N)$ には常用対数をとって 5 を引いた送金額を設定する．セルフループエッジにはエッジ特徴量を設定しない．

このようにして作成した取引関係グラフの例が図１である．この例にあるように，予測対象企業と継続した取引がある取引先は複数のエッジを持つのに対し，継続した取引が行われていない取引先はエッジが少なくなる．

図１：取引関係グラフの例



２．３．提案手法

取引関係グラフを入力として格付を予測するモデルを 4 層のニューラルネットワークで構築する．

まず 1 層目は Relational-GCN（[６]）を参考にして

下記に示す計算を行う． Relational-GCN との違いは正規化定数の代わりにエッジ特徴量を利用する点である．これによって，半期ごと・取引先ごとの送金額合計や受取額合計といった特徴量が内部的に生成される．

$$\boldsymbol{h}_i^{(1)} = \sigma\left(\sum_{j\in\mathcal{P}_{(i)}} W_p^{(0)} p_{j,i}\boldsymbol{h}_j^{(0)} + \sum_{k\in\mathcal{D}_{(i)}} W_d^{(0)} d_{k,i}\boldsymbol{h}_k^{(0)} + W_s^{(0)}\boldsymbol{h}_i^{(0)}\right)$$

ここで$\mathcal{P}_{(i)}$は i 番目のノードに向かって結ばれた payment エッジの始点ノード集合であり，$\mathcal{D}_{(i)}$は i 番目のノードに向かって結ばれた delivery エッジの始点ノード集合である．$W_p^{(0)}, W_d^{(0)}, W_s^{(0)}$は行数が隠れ層の次元数で列数が入力層の次元数（つまり，8）の重みパラメータ行列である．σは活性化関数であり，今回は ReLU 関数を用いた．

次に 2 層目は Graph Attention Network（[７]）を参考にして，下記に示す計算を行う．Graph Attention Network との違いはセルフループエッジとそれ以外のエッジで異なる重みパラメータを使用する点である．これによって，１層目で生成されたノードごとの特徴量を２層目に引き継ぎつつ，周辺ノードの特徴量を集約することが可能となる．

$$\boldsymbol{z}_i = W^{(1)}\boldsymbol{h}_i^{(1)}$$
$$\boldsymbol{s}_i = W_s^{(1)}\boldsymbol{h}_i^{(1)}$$
$$e_{i,j} = LeakyReLU\left(\boldsymbol{a}^{(1)^T}(\boldsymbol{z}_i||\boldsymbol{z}_j)\right)$$
$$f_i = LeakyReLU\left(\boldsymbol{b}^{(1)^T}\boldsymbol{s}_i\right)$$
$$\alpha_{i,j} = \frac{\exp(e_{i,j})}{\sum_{k\in\mathcal{P}_{(i)}}\exp(e_{i,k}) + \sum_{l\in\mathcal{D}_{(i)}}\exp(e_{i,l}) + \exp(f_i)}$$
$$\beta_i = \frac{\exp(f_i)}{\sum_{k\in\mathcal{P}_{(i)}}\exp(e_{i,k}) + \sum_{l\in\mathcal{D}_{(i)}}\exp(e_{i,l}) + \exp(f_i)}$$
$$\boldsymbol{h}_i^{(2)} = \sigma\left(\sum_{j\in\mathcal{P}_{(i)}}\alpha_{i,j}\boldsymbol{z}_j + \sum_{k\in\mathcal{D}_{(i)}}\alpha_{i,k}\boldsymbol{z}_k + \beta_i\boldsymbol{s}_i\right)$$

ここで$W^{(1)}, W_s^{(1)}$は行数・列数が隠れ層の次元数となっている重みパラメータ行列である．$\boldsymbol{a}^{(1)}$は隠れ層の次元数の２倍，$\boldsymbol{b}^{(1)}$は隠れ層の次元数を持った重みパラメータベクトルである．||はベクトルの連結を示す記号である．この層ではエッジ特徴量$p_{i,j}, d_{i,j}$を用いない．

3 層目は取引関係グラフの予測対象企業各半期のノード特徴量，内部取引のノード特徴量，取引先のノード特徴量の平均値，最大値を連結したベクトルを入力とした全結合層とする．

$$\boldsymbol{h}^{(2)} = (\boldsymbol{h}_1^{(2)}||\boldsymbol{h}_2^{(2)}||\boldsymbol{h}_3^{(2)}||\boldsymbol{h}_4^{(2)}||\boldsymbol{h}_5^{(2)}||\boldsymbol{h}_6^{(2)}||\boldsymbol{h}_7^{(2)} || \underset{i\in\mathcal{N}_p}{\text{mean}}\boldsymbol{h}_i^{(2)} || \underset{i\in\mathcal{N}_p}{\max}\boldsymbol{h}_i^{(2)})$$
$$\boldsymbol{h}^{(3)} = \sigma\left(W^{(2)}\boldsymbol{h}^{(2)} + \boldsymbol{b}^{(2)}\right)$$

ここで$\boldsymbol{h}_1^{(2)}$から$\boldsymbol{h}_6^{(2)}$は予測対象企業各半期のノード特徴量である．$\boldsymbol{h}_7^{(2)}$は内部取引ノード特徴量である．$\mathcal{N}_p$は取引先ノード集合である．$W^{(2)}$は行数が隠れ層の次元数の９倍で列数が隠れ層の次元数となっている重みパラメータ行列である．$\boldsymbol{b}^{(2)}$は隠れ層の次元数を持ったバイアスパラメータベクトルである．

4 層目は予測対象企業の格付が特定のカテゴリー以下である確率を出力する全結合層とする．

$$h^{(4)} = \text{sigmoid}\left(\boldsymbol{w}^{(3)^T}\boldsymbol{h}^{(3)} + b^{(3)}\right)$$

ここで$\boldsymbol{w}^{(3)}$は隠れ層の次元数を持った重みパラメータベクトルである．$b^{(3)}$はバイアスパラメータである．

プログラムの実装においては PyTorch（[８]）と Deep Graph Library（[９]）を用いた．各隠れ層の次元数を 16 とし，損失関数にクロスエントロピー誤差関数を用いて確率的勾配降下法にて各重みパラメータ・バイアスパラメータの最適化を行った．この際，学習率は 0.001，オプティマイザーは Adam，バッチサイズは 100，エポック数は 20 とした．

実装したソースコードは下記 URL で公開している．

https://github.com/marisakamozz/feature_extraction_from_bizgraph

２．４．比較手法

提案手法との比較を行うため，勾配ブースティング決定木で同様の予測モデルを作成した．

勾配ブースティング決定木では取引関係グラフをそのまま入力データとして使用できないため，取引関係グラフを元に集計特徴量とグラフ構造特徴量を作成し，入力データとした．

取引関係グラフの集計特徴量は，半期ごと・予測対象企業ごと・送金または受取ごとに取引先企業の件数，合計金額，最小金額，最大金額，平均金額，

中央金額，金額の分散，金額の標準偏差を求め，特徴量とした．また，送金合計金額と受取合計金額の差額を追加の特徴量とした．

また，取引関係グラフのグラフ構造特徴量として，グラフ全体のノード数・エッジ数，及び，次数１から次数６までのそれぞれのノード数を特徴量とした．ここで次数の少ないノードの数は継続した取引が行われていない取引先企業数を，次数の多いノードの数は継続して取引が行われている取引先企業数を意図した特徴量である．

勾配ブースティング決定木の実装には LightGBM（[１０]）を用いた．訓練データを学習データと評価データに８対２で分割した上で，学習データの５分割交差検証と Optuna（[１１]）で最適なハイパーパラメータを決定した．

３．テスト結果と考察

３．１．テスト結果

表２は提案手法と比較手法を使ったテストデータに対する予測結果の AUC スコアと AR 値を示している．

表２：提案手法・比較手法のスコア

	提案手法（GCN）	比較手法（LightGBM）
AUC	**0.794**	0.776
AR 値	**0.588**	0.552

３．２．テスト結果の考察

提案手法の AUC スコアは比較手法に比べて高い値となった．これは取引関係グラフのグラフ構造に格付を判別するにあたって有益な情報が内包されている事，そして，Graph Convolutional Network がそのグラフ構造からうまく特徴量を抽出できていることを示唆している．

また，比較手法に対して提案手法が優れている点として，比較手法においては取引関係グラフに内包されている特徴量をデータサイエンティストが手作業で設計・加工しなければならないのに対し，提案手法ではニューラルネットワークが格付の判別に有益な特徴量を自動的に抽出可能である事があげられる．

３．３．規模に対する精度・再現率の考察

表３・表４は取引関係グラフのノード数・エッジ数に対する提案手法の精度と再現率を示している．取引関係グラフにおけるノード数・エッジ数は取引先企業の数や取引件数を表しているため，ここから企業の規模を推し量ることが可能である．

表３：ノード数に対する精度・再現率

ノード数	特定カテゴリー以下の割合	精度	再現率
~10	0.31	0.67	0.79
~100	0.20	0.80	0.48
~1000	0.07	0.93	0.07
1001~	0.01	0.99	0.00

表４：エッジ数に対する精度・再現率

エッジ数	特定カテゴリー以下の割合	精度	再現率
~10	0.32	0.64	0.81
~100	0.25	0.75	0.58
~1000	0.11	0.89	0.17
1001~	0.03	0.97	0.01

ノード数・エッジ数ともに 100 以下の企業については精度・再現率ともに高い値となっているが，ノード数・エッジ数が増えるに従って精度は高くなるものの再現率は低くなっていく．これは大規模な企業ほど特定カテゴリー以下の格付の企業が少ないため，モデルは大規模な企業を優良企業と予測しがちになることが原因と考えられる．そのため，提案手法は比較的規模の小さな企業に対して有効なモデルであると言える．

４．まとめ

本研究では，金融機関が保有する送金明細データから取引関係グラフを構築し，Graph Convolutional Network を用いて企業の格付を予測する手法を提案した．提案手法は財務諸表に頼らず企業の返済能力を予測するための手法である．

提案手法を実際のデータに適用した結果，取引関係グラフには格付を予測するために必要な情報が内包されており，Graph Convolutional Network によって自動的にその情報を抽出できることが示唆された．また，提案手法は比較的規模の小さな企業に対して有効であることが示された．

提案手法の課題としては，取引関係グラフは予測

対象企業の直接の取引先に関する情報しか利用できない点があげられる．取引先のさらにその先にある取引先に関する情報も利用することで，サプライチェーン全体の構造を取り込むことができれば，より正確に格付を予測することが可能となる可能性がある．

参考文献

[１] Altman, E. I.: Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, The Journal of Finance, Vol. 23, No. 4, pp. 589–609 (1968)
[２] 高橋 久尚, 山下 智志:大規模データによるデフォルト確率の推定－中小企業信用リスク情報データベースを用いて－, 統計数理, Vol. 50, No. 2, pp. 241‐258 (2002)
[３] 本田 大悟，大古田 俊介，井實 康幸：入出金データを用いた企業デフォルト予測 ～機械学習手法の有効性比較評価～，第 48 回 2017 年度冬季 JAFEE 大会予稿集 (2018)
[４] 三浦 翔，井實 康幸，竹川 正浩：入出金情報を用いた信用リスク評価－機械学習による実証分析－，日本銀行ワーキングペーパーシリーズ，No.19-J-4 (2019)
[５] Thomas N. Kipf and Max Welling: Semi-supervised classification with graph convolutional networks, ICLR 2017 (2017)
[６] Michael Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov and Max Welling: Modeling Relational Data with Graph Convolutional Networks, arXiv:1703.06103 (2017)
[７] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Liò and Yoshua Bengio: Graph Attention Networks, ICLR208 (2018)
[８] PyTorch, https://pytorch.org/
[９] Deep Graph Library, https://www.dgl.ai/
[１０] LightGBM, https://github.com/microsoft/LightGBM
[１１] Optuna, https://optuna.org/

株式・債券・為替ごとの金融極性辞書の自動構築

Automatic Construction of Financial Sentiment Lexicons on Stock, Bond and Exchange

伊藤 友助[1]　酒井 浩之[1]　北島 良三[2]

末廣 徹[3]　稲垣 真太郎[3]　木村 柚里[3]

Tomosuke Ito[1], Hiroyuki Sakai[1], Ryozo Kitajima[2]

Toru Suehiro[3], Shintaro Inagaki[3], Yuri Kimura[3]

[1] 成蹊大学
[1] Seikei University
[2] 東京工芸大学
[2] Tokyo Polytechnic University
[3] みずほ証券株式会社
[3] Mizuho Securities Co., Ltd.

Abstract: In this paper, we propose a method of retrieving expressions and constructing financial sentiment lexicons on stock, bond and exchange. Retrieved phrases are distinctive words in each financial market and extensions of these words. For example, the sentiment of the word "increase" in the financial sentiment lexicons cannot be determined by merely looking at the word itself. However, we can determine sentiment by considering the relation between preceding and subsequent words, such as “increased consumption” and “increase in costs”. In addition, it is necessary that the expression such as “cause inflation” determines appropriate sentiment on stock, bond and exchange. We acquire automatically these expressions, and we construct appropriate financial sentiment lexicons by stock, bond and exchange.

１．はじめに

投資家は様々なデータを踏まえて投資判断を行うが，そのデータは膨大で多岐にわたるため，近年は膨大な金融情報を分析して投資判断を支援する技術が注目されている．その一例が「利上げ」のような金融に関連する語を収録した金融極性辞書である．

金融極性辞書について，これまでに株式に特化した「ネガティブ・ポジティブ」付与の研究が行われている．しかし，その極性辞書を債券などの他の金融市場に当てはめると極性が異なる場合が多い．これは同じ言葉でも金融市場によって極性が異なる場合があるためであり，例えば,「物価上昇」は株式市場においてはポジティブと判断するが，債券市場においてはネガティブと判断する必要がある．また，既存の金融極性辞書には単語（あるいは複合語）が収録されており,例えば,「増加」という語について，一つの極性が付与されている．しかし,「消費が増加」や「コストが増加」のように，語の前後関係まで広く捉えてみると，必ずしもその語に付与されている極性が正しいとは限らないことがわかる．

そこで本研究では，金融テキストを用いて，株式・債券・為替の金融市場ごとに表現にまで拡張した文字列での金融極性辞書を自動構築するための手法を提案する．本研究によって，辞書に収録するのに適した表現を獲得し,それらに極性を付与することで，各金融市場に特化した金融極性辞書の自動構築を実現する．

本研究では，金融極性辞書構築のためのデータとして，日経 QUICK 月次調査と日本経済新聞の記事を使用した．日経 QUICK 月次調査は，センチメントなどの定量的なデータが示されており，各金融市場のテキストデータを有するため，市場ごとの分析が可能である．

これまでに金融極性辞書についての研究は行われている[1][2][3][4]が，株式に特化した辞書が多く，辞書に収録されているのは単語（あるいは複合語）単位である．本研究では，株式・債券・為替の金融市

場ごとに語の前後関係まで拡張した表現を自動的に獲得する点が異なる．また，同一の表現でも金融市場によっては極性が異なる場合が考えられる．そのため，極性を付与するための基準となる情報を，例えば既存研究[3][4]のように株価のような１つの指標にすることはできず，各金融市場において適する指標を基準として，各表現に対して極性付与を行う必要がある．

２．提案手法の概要

本研究では，日経 QUICK 月次調査と日本経済新聞の記事を用いて，金融市場ごとに表現にまで拡張した文字列での金融極性辞書の自動構築を行う手法を提案する．金融極性辞書には各金融市場において特徴的な語を表現にまで拡張した文字列を格納する．提案手法の概要を以下と図１に示す．

Step１： 日経 QUICK 月次調査と日本経済新聞の記事を用いて，株式・債券・為替の金融市場ごとに金融極性辞書に収録するのに適した語の候補（以降，「金融市場特徴語候補」と定義）を抽出する．

Step２： 酒井らの手法[5]を用いて獲得した手がかり表現と抽出した金融市場特徴語候補をもとに，金融市場ごとに特徴的な語を表現にまで拡張した文字列（以降，「金融市場特徴表現」と定義）を獲得する．

Step３： 獲得した金融市場特徴表現に極性の付与を行い，金融極性辞書を構築する．

３．金融市場特徴語候補の抽出

3.1 金融市場特徴語候補抽出の概要

Step１： 日経 QUICK 月次調査に記載されている各金融市場担当者からのコメント文（以降，「コメント文」と定義）を抽出し，コメント文と日本経済新聞記事をもとに Doc2vec のモデルを金融市場ごとに生成する．

Step２： 生成した Doc2vec のモデルを用いて，金融市場ごとに類似した記事を日本経済新聞記事から抽出し，金融市場ごとにコメント文と合わせた文章データ（以降，「拡張文章データ」と定義）を作成する．

Step３： 酒井らの手法[5]を用いて，拡張文書データから金融市場ごとに共通頻出表現・手がかり表現を獲得する．

Step４： 拡張文書データをもとに，Word2vec のモデルを金融市場ごとに生成する．

Step５： 生成した Word2vec のモデルに獲得した共通頻出表現・手がかり表現を入力し，モデルから出力された語を金融市場特徴語候補として金融市場ごとに抽出する．

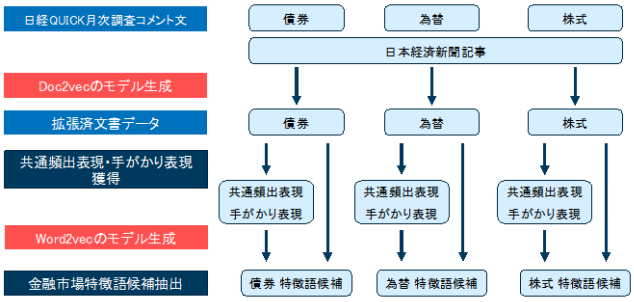


図１　金融市場特徴語候補抽出の概要

3.2 Doc2vec のモデル生成と類似記事抽出

日経 QUICK 月次調査に記載のコメント文を金融市場ごとに抽出する．しかし，抽出したデータが金融市場特徴語を獲得するデータ量として不足しているため，それらのテキストデータをもとに Doc2vec を用いて，日本経済新聞記事より金融市場ごとに類似する記事を抽出する．各金融市場においてそれぞれ Doc2vec のモデルを生成し，日経 QUICK 月次調査に記載のコメント文集合との類似度を求め，上位 50,000 記事を関連記事としてコメント文と合わせて拡張文書データとする．

3.3 共通頻出表現・手がかり表現獲得

酒井らの手法[5]を用いて，金融市場特徴語候補を抽出するための手がかりとなる表現を金融市場ごとの拡張文書データからそれぞれ自動的に獲得する．共通頻出表現・手がかり表現を獲得する手法について以下に述べる．なお，具体的な獲得手法については，文献[5]を参照されたい．

Step１： 少数の手がかり表現（今回は「が好調」「が不振」を用いる）を人手で与え，それに係る節を取得する．

Step２： 取得した節の集合から，その中で共通して頻繁に出現する表現（「売り上げ」など）を共通頻出表現として抽出する．

Step３： 共通頻出表現が係る節を取得し，その中から新たな手がかり表現を抽出する．

Step 4 : 獲得した手がかり表現から，それに係る節を取得する．

Step 5 : Step 2 から Step 4 を新たな手がかり表現と共通頻出表現が獲得されなくなる，もしくは，予め定めた回数まで繰り返す．

獲得した共通頻出表現・手がかり表現は形態素解析し，名詞と接頭詞，動詞の場合のみ結合した複合語とする．各金融市場における共通頻出表現・手がかり表現の例を以下に示す．

・株式

圧迫する，輸出関連株，下値不安，好調，収益率，不安材料，修正，経常赤字，落ち込む，維持する，一巡し，業績回復，押し下げ，借り入れ，弱まる

・債券

引き上げる，消費増税，高齢者，財政悪化，売却，先行し，縮小する，インフレ，債務残高，新興国，先行き不透明感，急減し，招き，浮上する，税収

・為替

後押しする，ユーロ高，期待感，政情不安，後退，投げ売り，再燃し，リスク回避姿勢，利上げ観測，中央銀行，増大する，下支えし，良好，財政支出

3.4 Word2vec のモデル生成と金融市場特徴語候補抽出

拡張文書データから金融市場特徴語候補を自動で取得するため，Word2vec を活用する．

Word2vec は単語の分散表現（ベクトル）を得る手法である．単語をベクトルで表現することで，単語間のコサイン類似度を計算することができ，意味のよく似た単語（類似度が高い単語）を抽出することができる．Word2vec のモデル生成時の学習データ（金融市場ごとの拡張文書データ）によって，同じ入力語であっても出力される類似度が高い単語が異なるため，金融市場における特徴的な語が出力されることを期待する．

なお，学習データは形態素解析したのち，連続する品詞や接頭詞の場合は結合し，複合語として作成した．金融市場ごとに先の学習データによって生成したモデルを用いて，前節で獲得した共通頻出表現・手がかり表現を入力とし，類似度が 0.5 以上の語を金融市場特徴語候補として抽出する．各金融市場における金融市場特徴語候補の例を以下に示す．

・株式

原価低減，半導体需要，自社株買い，可処分所得，危機感，軽自動車，買い戻し，失速懸念，薄らぐ，アジア向け輸出，潜在成長率，停滞感，くすぶり

・債券

ドル通貨オプション取引，銀行間市場，ベア，夢，中国経済，投資妙味，上向く，デフレ圧力，急伸，円高基調，低成長，買い進め，下げ止まる，顕著

・為替

原油依存度，日中韓，代わる，需給懸念，鎮める，中国人民銀行，衰退，米輸出入物価指数，叫ばれ，柔軟姿勢，取り崩す，実効為替レート，手詰まり感

3.5 金融市場特徴語候補の抽出結果

本手法により獲得した共通頻出表現・手がかり表現の数の詳細を表 1 に示す．また，抽出した各金融市場における金融市場特徴語候補の数の詳細を表 2 に示す．なお，「単独」はその金融市場にのみ存在する語の数を示しており，「総数」はその金融市場において得た語を合計した数を示している．

表 1　各金融市場の共通頻出表現・手がかり表現獲得数

金融市場	単独	総数
株式	170	1,230
債券	175	1,219
為替	72	1,096

表 2　各金融市場の金融市場特徴語候補の数

金融市場	単独	総数
株式	1,560	6,978
債券	1,494	7,009
為替	880	6,147

4．金融市場特徴表現の獲得

酒井らの手法[5]を用いて獲得した手がかり表現と抽出した金融市場特徴語候補をもとに，金融市場ごとに金融市場特徴表現を自動獲得する．表現獲得の概要を以下に示す．図 2 は株式市場における金融市場特徴表現の獲得の例を図示したものである．

4.1 金融市場特徴表現の獲得の概要

Step 1 : 獲得した共通頻出表現と手がかり表現の中から，手がかり表現のみを金融市場ごとに抽出する．(「が好調」「が不振」など)

Step 2 : 金融市場特徴語候補と抽出した手がかり表現を結合し，金融市場ごとに表現にまで拡張した文字列を生成する．

Step 3 : 生成した文字列が表現として適しているか判定を行い，文字列を金融市場特徴表現として獲得する．

・株式市場の例

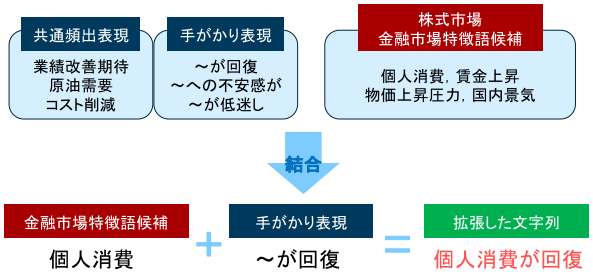


図２　株式市場における金融市場特徴表現獲得の例

4.2 表現にまで拡張した文字列の判定

本手法では，生成した文字列が表現として適しているか判定するために，拡張文書データを用いた．具体的には，金融市場ごとに金融市場特徴語候補と手がかり表現を結合して文字列を生成したのち，その文字列が金融市場ごとの拡張文書データに含まれているかの確認を行う．含まれていた場合，その文字列を金融市場特徴表現として獲得する．

4.3 表現獲得の結果

本手法によって獲得した金融市場特徴表現の一例を以下に示す．また，各金融市場における金融市場特徴表現獲得数の詳細を表３に示す．なお，表３には各金融市場において，拡張文書データに含まれていた回数ごとの表現の数，ならびに表現獲得数の総数を示している．

・株式

景況感指数が改善した，賃金上昇が続き，先行き不安が強い，市場が回復，輸出関連株が上昇し，個人消費への懸念が，景気減速懸念が強まった，需給が悪化して，物価上昇率が低迷，リスクが低く

・債券

米経済が回復し，均衡が拡大し，成長率が下がる，不透明感が強まる，赤字幅が縮小，政権が継続，信用リスクへの警戒感から，相場が伸び悩む，低金利政策が続く，言及がなかった，購入が必要

・為替

リスク回避姿勢が再燃，ドル高基調が高まる，財政赤字が続き，マイナス幅が縮小している，負債の多い，金利が低下，政策への期待感から，生産が増加，貿易赤字が続いて，警戒感が後退して

表３　各金融市場の金融市場特徴表現獲得数

金融市場	１回	＞２回	＞４回	総数
株式	53,148	20,522	11,343	86,759
債券	52,573	18,948	10,022	84,178
為替	43,647	16,176	8,751	70,386

4.4 金融市場特徴表現の評価

金融市場特徴表現についての評価は，金融市場ごとに獲得した表現のうち，拡張文書データに３回以上含まれている表現の中からランダムにそれぞれ 100 個抽出し，それらを人手によって正解を判定して行った．正解判定の基準は，ある表現について，その表現を獲得した金融市場において考えたときに，極性を付与することが可能かどうかで判断した．評価結果の精度を表４に示す．

表４　評価結果

金融市場	精度(%)
株式	88.0
債券	92.0
為替	80.0

５．金融市場特徴表現への極性付与

本手法により獲得した金融市場特徴表現に対して，対象としている市場に適した極性の付与を行う．しかし獲得した表現の数は多く，市場ごとに人手にて適切な極性を付与することは多大な労力を要する．そこで最小限の労力で極性付与のための学習データを自動的に生成し，深層学習により極性を付与することを試みた．極性付与の概要を以下に示す．

5.1 金融市場特徴表現への極性付与の概要

Step 1： 獲得した共通頻出表現と手がかり表現の中から，手がかり表現のみ（「が好調」「が不振」など）を金融市場ごとに抽出し， その中から手がかり表現単体で考えたときに，ネガティブまたはポジティブが明らかなものに対してその極性を人手で付与する．

Step 2： 金融市場ごとに拡張文書データ（類似する日本経済新聞記事上位 50,000 記事のみ）を文単位に分割し，文中に 4 章で獲得した金融市場特徴表現が含まれている文のみ抽出する．（以降，「テストデータ」と定義）

Step 3： テストデータの中から Step 1 で極性を付与した手がかり表現が含まれている文に対してその極性を付与し，学習データを作成する

Step 4： LSTM を用いた学習モデルを生成し，モデルをもとにテストデータに極性の付与を行う．

Step 5： テストデータに含まれる金融市場特徴表現に対して，テストデータである文に付与された極性をもとに極性の付与を行う．

5.2 手がかり表現への極性付与

3 章において，酒井らの手法[5]を用いて，金融市場特徴語候補を抽出するための手がかりとなる表現を金融市場ごとの拡張文書データからそれぞれ自動的に獲得を行った．その中から「が好調」「が不振」といった手がかり表現のみを金融市場ごとに抽出する．

金融市場ごとに抽出した手がかり表現の中から手がかり表現単体で考えた時に，その金融市場においてネガティブまたはポジティブが明らかなものに対して極性を人手で付与する．以下に，株式市場における手がかり表現への極性付与の例を図 3 に挙げる．

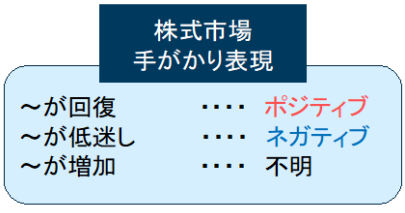


図 3　手がかり表現への極性付与の例

例のように，「～が回復」は，前にくる金融市場特徴語候補に関わらずポジティブであると判断できるため，手がかり表現に極性の付与を行う．同様に「～が低迷し」も明らかにネガティブの表現であると判断できる．しかし，「～が増加」については，前に出現する金融市場特徴語候補によって極性が変化するといえる．例えば，「リスクが増加」では株式市場においてネガティブと判断するが，「売り上げが増加」ではポジティブと判断する必要がある．このような手がかり表現に対しては，手がかり表現単体で考えたときに極性が明らかではないため，極性の付与は行わない．

5.3 学習データ・テストデータの作成

4 章にて獲得した金融市場特徴表現に極性を付与するための学習データを作成する．

まず， 3 章にて作成した金融市場ごとの拡張文書データの中から，金融市場ごとに類似する日本経済新聞記事 50,000 記事のみを抽出し，それらを文単位に分割する．記事単位であるデータを文単位に分割する理由としては，記事は全ての文において一つの極性が付与されているわけではなく，ネガティブ，ポジティブの文が混在して構成されており，記事への極性はそれらの文を総合的に判断して付与される．本研究の目的は，金融市場特徴表現への極性付与であり，文単位に極性を付与することによって，記事単位に極性を付与した場合よりも表現に対して正確な極性が付与できると考えるからである．

文単位に分割したのち， 4 章で獲得した金融市場特徴表現が文中に含まれている文のみを抽出し，これらをテストデータとする．さらに，テストデータの中から人手で極性付与した手がかり表現が含まれている文にはその極性を付与し，複数の手がかり表現が含まれる場合には最後に出現する手がかり表現の極性を文に付与し，極性が付与された文をまとめて学習データとする．

株式市場における学習データの例を以下に示す．なお，先頭の 1，0 は，学習データに付与された極性を示し，1 がポジティブ，0 がネガティブを表す．

1 000622-0220/382 家計の所得環境の緩やかな改善などを映し個人消費が堅調に推移するうえ、企業の設備投資も増加に転じる。
0 100929-1327/165 デフレが進行する中では、企業や家計の名目所得も大きく減少することから、税収が減少し、財政再建もおぼつかないだろう。

図 4 に本研究にて用いた LSTM＋多層パーセプトロン（MLP）による学習モデルを示す．

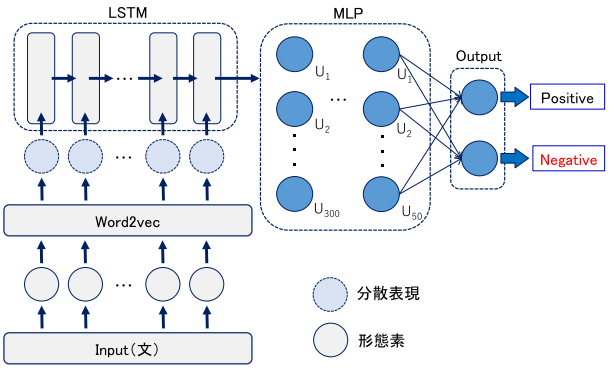


図４　LSTM＋多層パーセプトロン（MLP）

学習は以下の手順で行う．

Step 1： 学習データの文を形態素解析し，各語の分散表現（300 次元）を LSTM に入力する．

Step 2： 文頭から文末までのベクトルを多層パーセプトロン（MLP）の入力層（300 ユニット）に入力する．

Step 3： MLP の出力層は２ユニットとなっており，それぞれ，positive，negative の尤度を出力する．

MLP の中間層は 9 層（ユニット数 200 が 3 層，ユニット数 100 が 3 層，ユニット数 50 が 3 層）である．エポック数は株式市場においては 20，債券市場，為替市場においては 50 とし，活性化関数はランプ関数を使用した．また，分散表現を得るために word2vec を使用し，そのデータには日本経済新聞記事 10 年分を用いた．

本学習モデルは文の形態素の分散表現が文頭から文末まで推移していく過程が組み込まれたモデルであり，学習データにおける文の文脈を考慮した学習を行うことができる．

5.4 金融市場特徴表現への極性付与

作成した学習データを用いて上記の学習モデルにより，まずはテストデータに極性を付与する．テストデータである文のネガティブ，ポジティブそれぞれの尤度と，その文に含まれる金融市場特徴表現を獲得する．なお，尤度が高い方を優勢とし，その文に付与する極性とする．

次に，テストデータ各文のネガティブ，ポジティブの尤度とその文に含まれる金融市場特徴表現をもとに，金融市場特徴表現に対して極性の付与を行う．

ある文に含まれている金融市場特徴表現について，その文におけるネガティブ，ポジティブの尤度がその表現の尤度と等しいこととする．また，同じ金融市場特徴表現でも含まれる文によって尤度は異なるため，同じ表現のネガティブ，ポジティブの尤度の平均値をその金融市場特徴表現の尤度とする．これは一つの表現においてネガティブ，ポジティブの尤度を一意に決めるためである．

金融市場特徴表現の極性は，その表現のネガティブ，ポジティブの尤度が高い方を優勢とし，その極性とする．さらに，ネガティブ，ポジティブの尤度のどちらかが 0.7 以上の表現のみを抽出する．

各金融市場における金融市場特徴表現への極性付与数を表５に示す．また，株式市場における金融市場特徴表現への極性付与の例を表６に示す．

表５　各金融市場における金融市場特徴表現への極性付与数

金融市場	総数
株式	59,027
債券	61,590
為替	63,580

表６　株式市場における金融市場特徴表現への極性付与の例

極性	金融市場特徴表現	negative 尤度	positive 尤度
positive	復興需要への期待から	0.227	0.772
positive	消費心理が回復	0.015	0.984
positive	中間決算が好調だった	0.011	0.988
negative	警戒感が残る	0.723	0.276
negative	倒産が増加	0.829	0.171
negative	マイナス材料が増えて	0.957	0.044
negative	住宅バブルが後退して	0.784	0.216

６．評価

本手法についての評価は，極性の付与を行った金融市場特徴表現のうち，テストデータ内で出現する頻度が高いものから順に表現の先頭２文字が本評価のために抽出する表現と重複しない表現を 100 個抽出し，それらを人手によって正解を判定して行った．評価対象の表現について，先頭 2 文字が他の評価対象の表現と重複しない表現にする理由としては，表現の前半部分である金融市場特徴語候補が偏ることなく，多様な表現を評価するためである．なお，正解判定の基準はある表現について，その表現を獲得した金融市場において考えた時に，付与されている極性が正しいかどうかで判断した．株式市場，および，債券市場における評価結果を表７に示す．

表７　評価結果

金融市場	精度(%)
株式	75.0
債券	42.0

７．考察

金融市場特徴表現の獲得について，表４より，本手法による精度は株式 88.0%，債券 92.0%，為替 80.0%と高い精度を達成することができた．これは，金融市場特徴語候補と獲得した手がかり表現を結合し，その文字列が各金融市場に特化した拡張文書データに３回以上含まれている場合に金融市場特徴表現として評価対象としたためであると考える．獲得した表現は，実際に投資判断に役立てられている金融テキストデータに含まれている表現であるため，金融極性辞書に収録するのに適している語が多い．

しかし，表現が含まれていた回数を増やすことで精度が上がるとは言い切れず，例えば，株式の拡張文書データ内において 5 回出現する「状況がいつ」という表現が挙げられる．この表現は拡張文書内において,「状況がいつ好転するか不明である」などのような文の一部であると考える．このように，文書内において多く出現する表現であったとしても，極性を付与することができない表現が抽出されていることもあり，極性を付与するのに適していない表現の除去する手法が必要である．

次に獲得した金融市場特徴表現における極性付与について，本手法では尤度が 0.7 以上の表現のみを抽出して極性を付与したが，尤度の閾値を設定しない場合は，株式市場において精度が 46.0%という結果となり低かった．精度が低い理由としては，その金融市場特徴表現（本研究では，金融市場特徴語候補と手がかり表現を結合したもの）単体で考えた時に極性の判断ができないものが存在していることが考えられる．例えば，表８のような表現である．

表８　金融市場特徴表現への極性付与の例（尤度が低い場合）

極性	金融市場特徴表現	negative 尤度	positive 尤度
positive	拡大が続く	0.468	0.532
positive	地合いが続く	0.480	0.520
positive	注目度が高い	0.499	0.501
negative	動きが強まった	0.583	0.417
negative	傾向が続く	0.631	0.369
negative	ムードが強い	0.607	0.393

これらの表現は，その表現だけでは極性の判断が難しく，表現のさらに前にくる単語や表現によって極性を決めることができると考える．

評価において不正解と判定された表現について注目をしてみると，ネガティブとポジティブの尤度がそれぞれ 0.5 に近いものが多いことがわかった．どちらかの尤度が高い表現は本研究において極性がはっきりしていると言えるため，どちらかの尤度が 0.7 以上の表現のみを抽出し，再度評価を行った．株式市場において，評価に用いた金融市場特徴表現の例を表９に示す．

表９　金融市場特徴表現への極性付与の例（どちらかの尤度が高い場合）

極性	金融市場特徴表現	negative 尤度	positive 尤度
positive	効果が期待できる	0.016	0.984
positive	景気が回復	0.100	0.898
positive	材料が必要	0.266	0.734
negative	業績が悪化	0.929	0.071
negative	低迷が続く	0.845	0.155
negative	信用力の低い	0.722	0.278

これらの表現を用いて再度同様の方法で評価を行った結果が表７に示されている．表７より，株式市場において，精度が 75.0%と比較的良好な結果を得ることができた. このことから, 金融極性辞書には，表現のネガティブ，ポジティブのどちらかの尤度が 0.7 以上である金融市場特徴表現を格納することがふさわしいといえる．

また，表７より，債券市場における精度が 42.0%と低かった．これは，債券市場は株式市場と比較して, 極性の付与が複雑であることが理由だと考える．債券の場合は，一般的に景気が悪くなると安全資産であるため人気が上昇する．そのため表現の対象が景気であれば，ネガティブな表現でも極性をポジティブと判定しなくてはならない．例えば「景気が低迷」はポジティブと判定する必要がある．しかし，表現の対象が債券であれば，ネガティブな表現をそのままネガティブと判断する必要がある．例えば，「債券価格が低迷」という表現があげられる．そのような極性付与をするためには，学習データの作成時に手がかり表現の極性ではなく，その手がかり表現が対象としているものが「景気」なのか「債券」なのかを判定し，学習データを作成する必要があるが，本手法では学習データに含まれている手がかり表現の極性に従って極性を付与しているため，対応

できていなかった．

なお，為替市場の表現への極性付与ならびに金融極性辞書の構築も本手法と同様の方法で作成を行うことができると考えるが，対象が「円」である場合と「ドル」である場合で付与すべき極性が変化する．そのため，債券の場合と同様に学習データ作成時に対象についても考慮して作成する必要があると考える．

8．まとめ

本研究では，金融テキストを用いて，株式・債券・為替の金融市場ごとに表現にまで拡張した文字列での金融極性辞書を自動構築するための手法を提案した．

具体的には，日経 QUICK 月次調査と日本経済新聞の記事を用いて，はじめに株式・債券・為替の金融市場ごとに金融市場特徴語候補を抽出する．そして，獲得した手がかり表現と抽出した金融市場特徴語候補をもとに，金融市場ごとに金融市場特徴表現を獲得した．獲得した表現における評価の結果，精度は株式 88.0%，債券 92.0%，為替 80.0%と高い精度で表現を獲得することができた．

その後，獲得した金融市場特徴表現に極性の付与を行った．極性付与の結果を検証するためにテストデータ内で出現頻度が高く，かつ，評価対象の表現間で前半部分の重複がなく，ネガティブ，ポジティブのどちらかの尤度が 0.7 以上の表現を抽出した結果，株式市場において，精度が 75.0%と比較的良好な結果を得ることができた．

今後の課題としては，金融極性辞書に格納されている表現について，更に精度を向上させることである．特に債券市場，為替市場において表現の極性付与のための学習データ作成について別の手法も検討し，より最適な極性を付与することで，精度の向上を目指す必要がある．また，本研究では，各金融市場において金融市場特徴表現を格納している．しかし，単語または複合語ベースで最適な極性を付与することができる語も存在しており，それらに極性を付与し，表現とともに本研究の金融極性辞書に格納することでより実用性の高い極性辞書になることが期待できる．

参考文献

[1] Ito T., Sakaji H., Tsubouchi K., Izumi K., Yamashita T. (2018) 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.

[2] 伊藤諒，和泉潔，須田真太郎 : ネットワークの表現学習による金融専門極性辞書の構築，2017 年度人工知能学会全国大会(第 31 回)，2017.

[3] 五島圭一，高橋大志 : 株式価格情報を用いた金融極性辞書の作成，自然言語処理，Vol.24，No.5，pp547-577，2017.

[4] 関和広，柴本昌彦 : 銘柄固有の金融極性辞書の構築，第 18 回金融情報学研究会，2017.

[5] 酒井浩之，西沢裕子，松並祥吾，坂地泰紀 : 企業の決算短信 PDF からの業績要因の抽出，人工知能学会論文誌，Vol.30，No.1，pp.172-182，2015.

新興市場を対象とした市況情報の抽出

Extraction of market analysis information for emerging markets

神田 裕輝[1]　高野 海斗[1]　酒井 浩之[1]　北島 良三[2]　中川 慧[3]

Kaito Takano[1]　Yuki Kanda[1]　Hiroyuki Sakai[1]　Ryozo Kitajima[2]　Kei Nakagawa[3]

[1]成蹊大学
[1]Seikei University

[2]東京工芸大学
[2] Tokyo Polytechnic University

[3]野村アセットマネジメント株式会社
[3] Nomura Asset Management Co., Ltd.

Abstract: In this study, we propose a method to extract market information of emerging markets from newspaper articles. Here, we use Bloomberg articles as an information source to extract market information of emerging markets where is expanding in recent years. The extracted market information is useful as reference information when creating a market analysis report for the target emerging market. However, there are many articles that describe market conditions in large markets such as the Nikkei Stock Average and the Dow Jones Industrial Average, but there are few articles that mention one emerging market, such as India, China and Taiwan. It is only described in one part of an article that mentions several emerging markets. Therefore, in this study, we extract market information about one target emerging market from articles where information about market conditions of multiple emerging markets is mixed in one article. Furthermore, we select important sentences from the sentences extracted using the topic model and create a monthly report.

1　はじめに

近年，証券市場における個人投資家の比重が増加しており，個人投資家に対して投資判断の支援をおこなう技術の必要性が高まっている．その一例として，経済新聞記事をテキストマイニングの技術を用いて解析し，経済市場を分析する研究などが行われている[1][2][3]．また、投資判断のための情報源としてファンドの運用報告書に記載されるマーケットレポートが用いられる．マーケットレポートには世界経済や金融市場の動き等が記載されており、それらはファンドの運用担当者が、株価が大きく動いた日を調べ、その前後に起きたイベントを確認し、内容をまとめている．しかし、この作業は運用担当者にとって大きな負担となっている．したがって、経済テキストを機械的に解析し、マーケットレポートを自動で作成することができれば、ファンドの運用担当者の負担を減らすことができると考えられる．

本研究では，Bloomberg 記事を情報源として，近年市場が拡大しつつある新興市場を対象に分析を行い，市況情報を抽出する手法を提案する．本研究で抽出したい市況情報の例を以下に示す．

> ４日のインド株式市場では指標のＳ＆Ｐセンセックス指数が４営業日続落。不良債権が増えることへの懸念から銀行株が売られた。アクシス銀行はここ２日での下げが１年４カ月で最大となった。発電設備メーカーのバーラト重電機は１カ月ぶり大幅安。…..

関連研究として，経済テキストを用いて市況分析コメントを生成する研究がある．酒井らは日経新聞記事を用いて，日経平均株価について述べた記事やそれに関連のある記事を抽出し，そこから重要キーワード，重要な要因文を推定し，市況分析コメントを生成する手法を提案している[4][5]．しかし文献[4][5]の手法は日経平均について言及している複数の記事からレポートを生成するものであり，また，1 つの記事中で日経平均について言及している記事を対象としている．一方, Bloomberg においては１つの記事中で，例えばインド市場のような１つの新興市場について言及している記事はほぼなく，複数の新興市場の市況について言及している記事の１部に

記述があるだけである．そのため，新興市場の市況情報をまとめた市況分析レポートを作成するには，複数の新興市場の市況について言及している記事を探し，さらに，その中から対象としている新興市場の市況について述べている文を抽出し，それらをまとめる必要がある．従って，日経平均株価やダウ平均株価の市況分析レポートの作成に比べると多くの労力を必要とする．

そこで本研究では Bloomberg 記事を対象として，記事の 1 部のみに対象としている新興市場の記述がある状況下で，複数の記事より，対象新興市場についての有益な情報を抽出することを試みる．図 1 に，本研究において想定しているタスクの概要を示す．

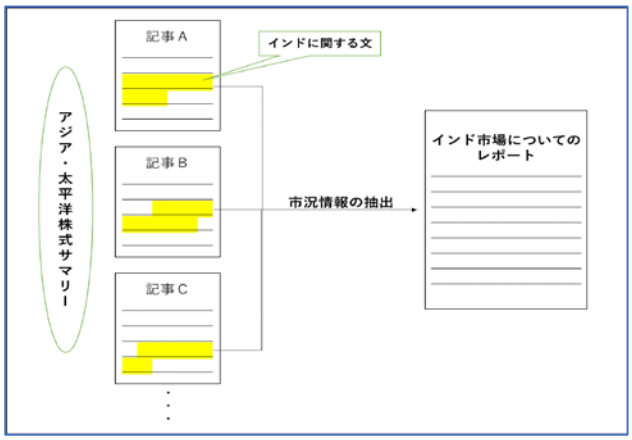


図 1　本研究のタスク概要

本研究では，対象とする新興市場をインド市場と定め，複数の新興市場の市況情報について記述されている「アジア・太平洋株式サマリー」からインド市場の市況情報のみを識別し，抽出する．

2　提案手法

本節では，Bloomberg 記事より対象の新興市場の株価に影響を与える情報を取得する手法の説明を行う．本研究では対象とする新興市場を「インド」と定めて研究を行った．

2.1　手法概要

複数の新興市場の市況情報についてまとめて記述されている「アジア・太平洋株式サマリー」には，インド市場以外にも中国市場や韓国市場，台湾市場といった複数の市場についての市況状況がある．そのため，文単位でインド市場の市況について述べている文であるかどうかの分類をする必要がある．手法の概要を以下に示す．

Step1：Bloomberg 記事から複数の新興市場について言及している記事（「アジア・太平洋株式サマリー」）を抽出．

Step2：抽出した記事から，最近傍法のための学習用データとして，対象としている新興市場（本研究ではインド市場）の市況について述べた文を抽出する．.

Step3：取得したデータから語の TF･IDF 値を計算し，語を要素，TF･IDF 値を要素値とした単語ベクトルを生成する

Step4: 学習用データとテスト用データとの文間類似度を求め，最近傍法にてインド市場の市況情報について述べている文であるかどうかを分類する

Step5：抽出した文を時系列順に並べ，市況情報をまとめる．

2.2　新興市場に関する記事の抽出

本研究では，Bloomberg 記事の 2015 年 1 月から 2017 年 9 月の約 3 年分の記事を用い，使用する記事数は新規記事 4,338,113 記事である．

新興市場（インド市場）について言及している記事を抽出するために，記事本文に対象とする国名と経済や市場に関する単語を含む記事を抽出した．内容を確認したところ，その多くが「アジア・太平洋株式サマリー」の記事であった．そのため，特定の市場（インド市場）に限定したレポートを生成するには，この記事から特定の市場の市況情報について言及した文のみを抽出する必要がある．そこで，この記事中の“インド株式市況”の前後に注目して，重要な文の抽出を試みる．なお，抽出された「アジア・太平洋株式サマリー」は 712 記事である．

2.3　学習データの作成

前節で抽出された「アジア･太平洋株式サマリー」712 記事を用いて，インド市場の市況情報に関する文を抽出するために，最近傍法により文をインド市場に関する文かそうでないかを判定する．まず，学習用データとテスト用データを作成する．ここで，“インド株式市況”を含む文とその次の文は，インドについて言及していると想定し学習用データの正例とする．この時に明らかに他の国の市場について言及されている場合は正例データから除くこととする．一方で，“インド株式市況”を含む文のそれ以前の文はインド以外について言及していると想定し学習用データの負例とする．以下に学習データの例を示す．

・正例

・【インド株式市況】 6 日のインド株式相場は 6 営業日続落。タタ・モーターズの利益が市場予想に届かず、企業業績に照らした株価水準に警戒感が広がった。
・【インド株式市況】21 日のインド株式相場は反落。タタ・スチールの四半期赤字が予想を上回る規模だったことから、金属株が売られた。

・負例

・【中国・香港株式市況】中国本土の株式相場は 3 営業日続落。
・原油安で利益が押し上げられ、消費者の需要も高まるとの見方が材料視された。
・ペトロチャイナ（601857CH）は 6 %高と、4 年ぶりの高値を付け、上海総合指数の上昇に大きく寄与した。

テスト用データは正例とした 2 文以降の文とする．この時抽出した文は，正例が 1403 文，負例が 9696 文，テスト用データが 6896 文である．また，以降の市況情報の抽出の計算において文のつながりを考慮し，正例の 2 文を繋げ 1 つの文として想定することにより，正例を 706 文とする．

2.4 市況情報の抽出

単語の出現頻度に基づいた重みを計算するために式(1)より TF･IDF 値を算出する．使用する単語は半角の単語を除いた 2 語以上の単語とする．

$$TF \cdot IDF(t,d) = TF(t,d) \cdot \log\left(\frac{|N|}{df(t,N)}\right) \quad (1)$$

N：学習用データの文の総数（今回は 10402 文）
$TF(t,d)$：文$d \in N$において，単語 t の出現頻度
$df(t,N)$：文集合 N のうち単語 t を含む文の数

算出した TF･IDF 値をベクトルの要素値として，式(2)よりベクトル間の余弦を求め，文間類似度 sim を計算する．

$$\mathrm{sim}(V_d, V_t) = \frac{V_d \cdot V_t}{\|V_d\| \cdot \|V_t\|} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}} \quad (2)$$

ここで，V_d, V_tはそれぞれ学習用データ，テスト用データに含まれる単語を要素とした単語ベクトル（$V_d = \{x_1, x_2, \cdots, x_n\}$, $V_t = \{y_1, y_2, \cdots, y_n\}$），$x_i, y_i$はそれぞれ学習用データ，テスト用データに含まれる単語の TF･IDF 値を表す．

テストデータの文と学習データとの文間類似度を求め，学習データにおける最も類似度が高い文が正例であったのであれば，その文をインド市場の市況情報についての文であるとして抽出する．全テスト用データ 6896 文を最近傍法で判定した結果，インド市場の市況情報に関わる文として抽出できたのは 2132 文であった．以下に，抽出された文の一部を示す．

・世界最大の製油所を所有するリライアンス・インダストリーズは 10 カ月ぶり安値を付けた。
・インド石油ガス公社（ＯＮＧＣ）は構成銘柄の中で最もきつい値下がり。
・インド最大の民間銀行、ＩＣＩＣＩ銀行は続落。
・タタ・スチールが大きく下げ、金属株指数は 3 週間ぶりの低水準に落ち込んだ。

2.5 評価

テスト用データの 100 文をインドに関する文かそうでないかを最近傍法によって判定した．テスト用データ 100 文を，正例・負例を含めた学習用データ 10402 文とのそれぞれの文間類似度を算出し，その中で最も値が大きい文と同じ分類にすることとした．ここで，文中に“インド”を含むものは，文間類似度の値によらずインド市場に関係すると判定する．分類の内容としてインド市場に関係すると判定した場合 1，インド市場とは関係ないと判定した場合は-1 とする．また，それらの分類分けが正しいかどうかを人手にて○×で評価した．表 1 に判定結果の例を示す．

表 1. 文間類似度の判定結果の例

	テストデータの文	分類	評価
①	英高級車メーカーのジャガー・ランドローバーを傘下に置くタタ・モーターズは３日続伸	1	○
②	政府の１日声明によれば，道路建設に充てる資金捻出のため燃料税がリットル当たり２ルピー引き上げられた	−1	×

③	【韓国株式市況】韓国総合株価指数は前日比1.1%高の1904.65	−1	○
④	銀行株と公益事業株の上げが目立った	1	×

この評価より，最近傍法による全体の精度は 76%であることがわかった．

2.6 レポートの作成

最近傍法により抽出した 2132 文と，その文を含む記事における正例の文を繋げ，それを時系列順に並べて，月ごとのインド市場の市況情報のレポートを作成する．以下に例として抽出した 2015 年 2 月の市況情報の一部を示す．正例データの文と，最近傍法により抽出された市況情報の文を組み合わせることによって，繋がりのあるわかりやすい文章になっている．

２日のインド株式市場では指標のＳ＆Ｐセンセックス指数が前週末に続き下落。消費関連銘柄などが売られた。インド準備銀行（中央銀行）は３日に金融政策決定会合を開く。たばこのＩＴＣは続落。センセックス指数は前週末比 0.2%安の 20122.27 で終了。インド中銀の３日の金融政策決定を前にブルームバーグ・ニュースがまとめたエコノミスト調査では、41 人中 31 人が 7.75%での政策金利据え置きを予想。４日のインド株式市場では指標のＳ＆Ｐセンセックス指数が４営業日続落。不良債権が増えることへの懸念から銀行株が売られた。アクシス銀行はここ２日での下げが１年４カ月で最大となった。発電設備メーカーのバーラト重電機は１カ月ぶり大幅安。.....

3 考察

最近傍法による分類であるが，全体の精度は 76%であった．まず，表 1 の①のようにインドに関する文が正しく判定されたものは 47 文中 24 文あり，これらの多くは，文中にインドに関する固有名詞を含んでいたため，正しく判定されたと考えられる．

表 1 の②のようにインドに関わる文を誤って判定したものは 23 文あった．この時，表 1 の例のようにどの国にも共通する単語が多くみられ，インドに関する固有名詞が含まれていたが負例と同じ分類に判定された文もあったため，df 値の取得方法を改める必要があると考えられる．

表 1 の③のようにインドとは関係のない文で正しく判定されたものは 53 文中 52 文あり，これらは例のようにインド以外の国名が含まれることや，インドに関する固有名詞が含まれていないことが多かったため正しく判定できたのだと考えられる．

表 1 の④のようにインドとは関係のない文を誤って判定したものは 1 文あった．これらはどの国にも関係しうる文であったため，判定が困難であったと考えられる．

また今回，最近傍法との比較のため深層学習による市況情報の抽出も試みた．使用した学習用データとテスト用データは，最近傍法の学習データの作成方法と同様である．ただし，各記事の正例の文を繋げ，正例が 706 文，負例が 9696 文，テスト用データが 6896 文とした．深層学習モデルは，Long short-term memory (LSTM) と Multilayer perceptron (MLP) を組み合わせ，図 2 で示すモデルを用いた．

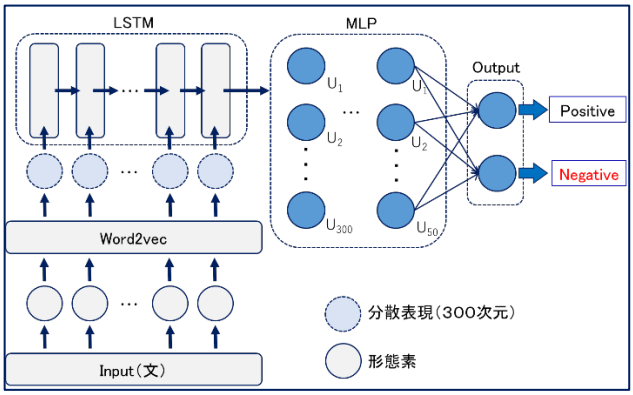


図 2　深層学習モデル

処理の概要であるが，はじめに学習データの文を形態素解析し，Word2vec により形態素の分散表現を得る．それを LSTM に入力し，出力されたベクトルを MLP の入力としてモデルを学習する．MLP の中間層は，ノード数 300 が 3 層，ノード数 150 が 3 層，ノード数 50 が 3 層の計 9 層とした．出力層は 2 要素である．また，epoch 数は 3，中間層の活性化関数を ReLU とし，出力層の活性化関数をシグモイド関数とした．なお，分散表現を得るための Word2vec の学習データとして，10 年分の経済新聞記事を使用した．

全テスト用データ 6896 文を判定した結果，インド市場の市況情報に関わる文として抽出されたのは 315 文でり，また，その精度は 72%であった．この結果は最近傍法の結果である 83%と比べて低いものであり，Bloomberg 記事を対象とした本研究においては最近傍法が有用である結果であった．

さらに手法比較にくわえて，本提案手法を他の新興国市場を対象としても正しく市況情報抽出が可能かどうかを検証するために，対象を「中国・香港市

場」として追加実験を行った．その結果抽出された例を以下に示す．

・上海総合指数は前営業日に当たる昨年 12 月 31 日と比べ 3.6%高の 3350.52 と、終値ベースで 09 年 8 月 6 日以来の高値を付けた
・ペトロチャイナ（601857CH）は 6 %高と、4 年ぶりの高値を付け、上海総合指数の上昇に大きく寄与した

テスト用データ 100 文を最近傍法により判定し，その結果を人手で評価したところ，精度は 77%であった．この結果から，本手法では高い精度を維持しながら他の新興国を対象としても同様の処理が可能であり，汎用性の高い手法と言える．

4　トピックモデルによるレポートの生成

2 章では最近傍法によって抽出した文と同記事の正例に当たる文を繋げることによって，市況情報をまとめたレポートを作成した．ただし，抽出した文は 2132 文あり，レポートを作成にするには文量が多い．そのため，レポートに採用する文を最近傍法によって抽出した文の中から選別する必要がある．そこで本節では，重要な文の選別方法とレポートの作成方法について説明する．

4.1　トピックモデルの概要

本研究では Python のライブラリのうち gensim の LDA を用いてトピックモデルを生成する．図 3 に本手法の流れを示す．

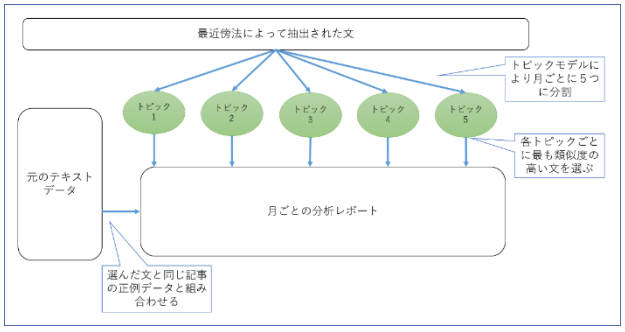


図 3　トピックモデルを用いたレポート生成

抽出した文から重要な文を選別する方法として，まず最近傍法の際に用いた学習用データの文をデータセットとして，gensim によってトピックモデルを生成する．この時のモデル生成では，数字を除く 2 語以上の名詞，動詞，形容詞を対象とし，トピック数を 5 に設定した．これによって似た単語を持つ文で 5 つに分割できる．表 2 に生成されたトピックの例を示す．そして生成されたモデルを最近傍法によって抽出した文に月ごとに適応することで，抽出した文を 5 つのトピックに分割する．次に，各トピックで最近傍法の際に求めた文間類似度が最も高い文をレポートに採用する．つまり，各月で 5 文選ばれることになる．そして，選んだ文と同じ記事の正例データの 1 文（記事のうちインドに言及している最初の文）と繋げ，最大合計 10 文のレポートを作成する．

表 2．生成されたトピックの例

トピック 1		トピック 2		トピック 3	
単語	生起確率	単語	生起確率	単語	生起確率
銘柄	0.028	指数	0.120	株式	0.131
下げ	0.026	中国	0.068	市況	0.052
発表	0.022	セン	0.056	インド	0.052
なっ	0.022	ハン	0.054	相場	0.041
こと	0.018	本土	0.051	香港	0.039
利益	0.016	前日	0.042	市場	0.038

4.2　トピックモデルの評価

最近傍法によって抽出した 2132 文を月ごとにモデルにかけ，5 つのトピックに分割した後，レポートを作成した．表 3 に例として 2015 年 3 月のトピックモデルによる結果を示す．

表 3　2015 年 3 月の各トピックにおける単語の生起確率

トピック 1		トピック 2	
単語	生起確率	単語	生起確率
会社	0.006	インド	0.010
ぶり	0.006	ルピー	0.008
最大	0.005	銀行	0.005
週間	0.005	ドル	0.005
インド	0.005	高値	0.005
セン	0.004	上昇	0.005
大幅	0.004	付け	0.004
セックス	0.004	バラード	0.004
トピック 3		トピック 4	
単語	生起確率	単語	生起確率
セックス	0.017	ぶり	0.009
セン	0.016	大幅	0.004

前日	0.016	セン	0.004
終了	0.010	セックス	0.004
引け	0.005	終了	0.004
銀行	0.005	週間	0.004
ぶり	0.004	HDFC	0.004
インド	0.004	住宅	0.003

トピック 5

単語	生起確率
メーカー	0.004
続落	0.004
最大手	0.003
営業	0.003
ぶり	0.003
なっ	0.003
電機	0.003
発電	0.003

・トピック 1 に属する文

・指標のＳ＆Ｐ・ＢＳＥセンセックスは週間ベースで４週続伸となった
・世界最大の製油所を所有するリライアンス・インダストリーズは 1.4%安
・国内最大のエンジニアリング会社ラーセン・アンド・トゥブロは２週間ぶりの大幅下落、インドステイト銀行は５営業日続落となった
・インド最大のエネルギー探査会社、インド石油ガス公社（ＯＮＧＣ）は５カ月ぶりの大幅上昇

・トピック 2 に属する文

・ルピー建て 2024 年７月償還債の利回りは２ベーシスポイント（ｂｐ、１ｂｐ＝0.01%）上昇の 7.74%
・４日にはインド準備銀行（中央銀行）による緊急利下げを受けて、初めて３万台を付けた
・通貨ルピーとインド国債も値下がりした
・構成銘柄の中で発電設備のバーラト重電機の下落率が最大だった

・トピック 3 に属する文

・インド株の指標であるＳ＆Ｐ・ＢＳＥセンセックスは前週末比 0.3%高の 29459.14 で終了
・センセックスは前日比 0.7%安の 29380.73 で終了
・センセックスは前日比 0.2%高の 29448.95 で終了した
・指標のＳ＆Ｐ・ＢＳＥセンセックスは前営業日比 2.1%安の 28844.78 で終了

・トピック 4 に属する文

・住宅金融でインド最大手のハウジング・デベロップメント・ファイナンス（ＨＤＦＣ）は２カ月ぶりの大幅下落
・英高級車メーカーのジャガー・ランドローバーを傘下に置くタタ・モーターズと、エンジニアリング会社ラーセン・アンド・トゥブロも買われた
・住宅金融のハウジング・デベロップメント（ＨＤＦＣ）は３日ぶりに値上がり
・英高級車メーカーのジャガー・ランドローバーを傘下に置くタタ・モーターズは６週間ぶりの大幅下落

・トピック 5 に属する文

・日用品大手ヒンドゥスタン・ユニリーバは１カ月ぶり高値に上昇
・発電設備メーカーのバーラト重電機は３営業日ぶりに下落
・一方、契約者数でインド最大手の携帯電話サービス会社ブハルティ・エアテルは続伸した
・携帯電話サービス企業ブハルティ・エアテルを中心に電気通信株が下げた

単語の生起確率と各トピックの文を見ると，トピックごとに特徴があることが分かる．トピック 1 では生起確率の上位に“最大”があり，トピック 1 の文を見るとインドの大手企業について言及している文が多い傾向にある．同様にトピック 2 では通貨や銀行について，トピック 3 ではインド株の指標について，トピック 4 では自動車や住宅金融メーカーについて，トピック 5 ではその他メーカーについて言及している傾向にあった．このようにトピックモデルによって抽出した文を 5 つのトピックに分割した後に，各トピックから 1 文をレポートに採用する文として選ぶことによって，情報が重複しないレポートを生成できる．以下に作成したレポートを示す．

・2015 年 3 月のインドの市況情報のレポート

２日のインド金融市場では株価指数が３営業日続伸。インド株の指標であるＳ＆Ｐ・ＢＳＥセンセックスは前週末比 0.3%高の 29459.14 で終了。
５日のインド株式相場は上昇。指標のＳ＆Ｐ・ＢＳＥセンセックスは週間ベースで４週続伸となった。４日にはインド準備銀行（中央銀行）による緊急利下げを受けて、初めて３万台を付けた。
10 日のインド株式市場では指標のＳ＆Ｐ・ＢＳＥセンセックスが続落し、１カ月ぶりの安値を付けた。住宅金融でインド最大手のハウジング・デベロップ

メント・ファイナンス（ＨＤＦＣ）は２カ月ぶりの大幅下落。
インド株式市場では指標のＳ＆Ｐ・ＢＳＥセンセックスが下落。発電設備メーカーのバーラト重電機は３営業日ぶりに下落。

下線部が選別した文を示す．選んだ文と同記事の正例の 1 文と繋げることによって読みやすく分かりやすい文章になっている．また，抽出した文は時系列順に表示し、同じ記事の文を抽出した場合はそれらを繋げることで違和感のない内容のレポートになっている．

また，各トピックから文を抽出せずに 1 つのトピックからいくつかの文を抽出すれば，1 つのトピックに関するレポートを生成することもできる．例えば，金融に関するレポートを生成したければ，金融に関するトピックから文を抽出することで可能になると考えられる．

しかし，トピック間で明確に分割できていない文もあり，似た文が異なるトピックに属していることも多い．これを改善するために，トピックモデルを生成した際のデータを見直すことや，特徴のある単語には重み付けする必要があると考えられる．

4 まとめ

本稿では Bloomberg 記事をもとに，複数の記事からインド市場のような 1 つの新興市場の市況情報を抽出する手法について述べた．本手法では新興市場のようなニッチなものは記事の一部に記述されていることに注意し，最近傍法によって必要な情報を抽出した．そしてトピックモデルによって抽出した文から重要な文を選別し，市況情報をまとめたレポートを作成した．本手法の精度は 76%となり，比較手法として行った深層学習の手法よりも高い精度を得ることができた．また，他の新興市場を対象として市況情報を抽出することによって，本手法の有用性・汎用性を確かめられた．

今後の課題として，最近傍法とトピックモデルの精度を向上することや，文章にする上で要因文を推定することが考えられる．

参考文献

[1] 蔵本貴久，和泉潔，吉村忍，石田智也，中島啓浩，松井藤五郎，吉田稔，中川裕志: 新聞記事のテキストマイニングによる長期市場動向の分析，人工知能学会論文誌, Vol. 28, No. 3, pp. 291-296, 2013.

[2] 和泉潔，後藤卓，松井藤五郎: 経済テキスト情報を用いた長期的な市場動向推定，情報処理学会論文誌, Vol. 52, No. 12, pp. 3309-3315, 2011.

[3] 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・業績結果の抽出，人工知能学会論文誌, Vol. 34, No. 5, pp1-22, 2019.

[4] 酒井浩之，坂地泰紀，和泉潔，松井藤五郎，入江圭太郎: 関連記事を用いた市況分析コメントの自動生成，第 22 回 金融情報学研究会, pp. 61-66, 2019.

[5] 酒井浩之，坂地泰紀，和泉潔，松井藤五郎，入江圭太郎: 経済テキストからの市況分析コメントの自動生成，第 20 回 金融情報学研究会, pp.44-49, 2018.

テキストマイニングを利用した テーマに関連する上場企業検索ツールの開発

Development of Search Tool for Listed Company Related to Themes Using Text Mining

平野正徳 [1*] 坂地泰紀 [1] 木村笙子 [2] 和泉潔 [1] 松島裕康 [1] 長尾慎太郎 [2] 加藤惇雄 [3]
Masanori HIRANO[1] Hiroki SAKAJI[1] Shoko KIMURA[2]
Kiyoshi IZUMI[1] Hiroyasu MATSUSHIMA[1] Shintaro NAGAO[2] Atsuo KATO[3]

[1] 東京大学 大学院工学系研究科
[1] School of Engineering, The University of Tokyo
[2] 大和証券投資信託委託株式会社 調査部
[2] Research Department, Daiwa Asset Management Co. Ltd.
[3] 株式会社大和総研 フロンティアテクノロジー本部
[3] Frontier Technologies Research & Consulting Deptartment, Daiwa Institute of Research Ltd.

Abstract: We propose a scheme for selecting stocks related to a theme. This scheme was designed to support fund managers who are building themed mutual funds. Our scheme is a type of natural language processing method and based on words extracted according to their similarity to a theme using word2vec and our unique similarity based on co-occurrence in company information. We used data including investor relations and official websites as company information data. We also conducted several other experiments, including hyperparameter tuning, in our scheme.

1 はじめに

近年，日本において，個人投資家の数が増えている．個人投資家の増加に対応し，個人投資家をサポートするさまざまな技術が開発されている．ニュース記事の中から自動的に景気動向を示す表現を抜き出す手法 [Sakaji 08] や，上場企業が公開する，共通形式の決算速報である決算短信の中から自動的に業績関連文を取り出す手法 [Kitamori 17]，金融テキストのセンチメントを可視化する解釈可能なテキストマイニング [Ito 18b, Ito 18a] などがあげられる．

日本の個人投資家の間では，特に投資信託と呼ばれる，投資家から集めた資金で投資信託会社などが資金運用をして，成果を投資家に還元する金融商品が人気がある．似たような金融商品の一種として，世界中で取引されている Exchange Traded Funds (ETF) が有名であるが，投資信託と ETF は大きく異なる．ETF は一般に，それぞれの ETF を構成する構成銘柄の比率は固定されている．一方で，投資信託は，投資信託を管理している金融機関や投資信託会社などが，動的に構成銘柄やその構成比率を変更させている．

投資信託の中には，テーマ型投資信託 (テーマ型ファンド) と呼ばれる投資信託が存在する．特にテーマ型投資信託は日本の投資家に人気があり，人工知能（AI），ロボティクス，健康など特定の分野を投資先として選んだものとなっている．これらのテーマ型投資信託は，テーマに関連する株式への投資を通じて，テーマの盛り上がりに応じた収益を得ることを目指している．

投資信託会社は，さまざまな種類のファンドを立ち上げ，管理・運用をしている．特に，投資家を引き付けるためには，さまざまな種類のテーマ型ファンドを売り出すことが重要であるとともに，テーマの盛り上がりに対して，適切なタイミングでファンドを売り出していくことが重要である．

テーマ型ファンドの開発には多くの課題がある．テーマ型投資信託を作成するために，投資信託会社は，売れそうな投資信託のテーマを探し，さらに，テーマに

*連絡先：東京大学 大学院工学系研究科 和泉研究室
〒 113-8656 東京都文京区本郷 7-3-1
E-mail: hirano@g.ecc.u-tokyo.ac.jp
HP: https://mhirano.jp/

†本稿は，[Hirano 18, Hirano 19a, Hirano 19b, Hirano 19c] の内容を再構成したものです．

‡特許登録済 (特許第 6596565 号)

§本論文の内容や意見は執筆者個人に属し，いかなる組織の公式見解を示すものではありません．また，特定の商品についての投資の勧誘及び売買の推奨を目的としたものではありません．

関連する銘柄を探さなければならない．加えて，その中から運用において優秀な銘柄を選択してポートフォリオを作成しなければならない．関連する銘柄を探すことは，投資信託を運用するファンドマネージャーにとって，時間的に多大な負担であり，さらに，人手で行うと，本当はテーマに関連のある銘柄を見落としてしまう可能性もある．また，ファンドマネージャーは必ずしもテーマに関して詳しいわけではなく，そういったテーマを取り扱う場合には，事前のリサーチを行うことで対応をしているが，グローバルな投資信託，つまり，海外の銘柄を含むようなファンドの場合，全ての銘柄をカバーすることはほぼ不可能である．そこで，ファンドマネージャーの負担を軽減し，テーマに関連する銘柄を見落としなく抽出するためには，この手順を自動化することが必要である．

近年，様々な技術が発展してきている．特に，テキストマイニングやビッグデータなどの技術は，この分野に応用可能であると考えられる．本論文においては，テキストマイニングに着目して，自然言語処理を用いた，関連銘柄抽出手法を提案する．

2 先行研究

テーマ型投資信託における，ファンドマネージャーをサポートするようなシステムの構築は，現在の日本市場においてはまだまだ先進的な取り組みであり，新規性の高いタスクである．そのため，ここでは，関連する分野における先行研究について述べる．

まず，金融にテキストマイニングを利用している研究をここでいくつか取り上げる．Koppel らは企業の株価のデータを利用して，企業のニュースが良いニュースなのか悪いニュースなのかを分類するテキストマイニングの手法を提案している [Koppel 06]．Low らは “semantic expectation-based knowledge extraction methodology (SEKE)” というニュースなどのテキストから因果関係を抽出する手法を提案しており [Low 01]，単語の概念に関するシソーラスとして，WordNet [Fellbaum 98] を使用している．Schumaker らは金融に関するニュース記事をもとに，株価の予測を行うということを機械学習を利用したアプローチで行なって，その有効性を確かめた [Schumaker 09]．さらに，Ito らは “gradient interpretable neural networks (GINN)” という，金融テキストのセンチメントを可視化するニューラルネットを利用したモデルを提案しており，テキストマイニングにおいて，解釈可能性を追求している [Ito 18b, Ito 18a]．Milea らは，MSCI EURO index の変動予測 (上昇，下降，変化なし) の予測を，European Central Bank (ECB) の発表する文書の文脈から予測する手法を提案している [Milea 10]．Xing らは，アメリカの株式市場における，企業間のつながりを株価の変動の特徴に加えて，テキストマイニングを利用することでより良く抜き出す手法を利用し，アセットアロケーションタスクに組み込むことで，企業間の関係を正確に把握し，ポートフォリオを最適化する手法を提案している [Xing 18]．

次にテキストマイニングを金融の分野に応用している日本語をターゲットとした研究について述べる．Sakai らは日本語のニュースのうち，業績について書かれた文章の中の因果関係を抜き出す手法を提案しており，ブートストラッピング的な手法により，自動的に因果関係を抜き出すための因果を示す手がかり表現を抜き出した [Sakai 07]．Sakaji らはニュース記事から統計的な手法により，自動的に景気動向を示す表現を抜き出す手法を提案した [Sakaji 08]．さらに，Sakaji らは決算短信の中から自動的に稀な因果関係を抜き出してくる手法を提案した [Sakaji 17]．Kitamori らは決算短信の中から業績に関する文章をニューラルネットワークモデルを利用して取り出し，分類する半教師あり学習の手法を提案している [Kitamori 17]．

3 手法

本章では，提案手法について説明する．提案手法はファンドのテーマにおける関連銘柄を選ぶというタスクに対して，ファンドマネージャーをサポートするための手法である．手法概要は図 1 の通りである．まず，テーマの単語をインプットとして入力すると，Word2vec を利用した類似度と企業情報における単語の共起に基づいた類似度を計算し，それらを合わせて最終的な単語の類似度を計算する．単語の類似度を使用して，元のテーマの単語に関連する単語を選び，その単語を含む文を企業情報の中から抽出する．さらに，抽出された文章内に含まれる最終的な関連単語の出現回数とその類似度を加味して企業の類似度を計算する．その結果に基づき，企業の関連度ランキングを作成し，アウトプットとする．この際，抽出の根拠文として，抽出された最終的な関連単語を含む企業情報の文を一緒にアウトプットする．

3.1 Word2vec モデルの構築とそのモデルを利用した類似度計算

まず，テーマの単語の入力を受けた後，Word2vec を利用した類似単語の抽出と類似度計算を行う．Word2vec とは，Mikolov ら [Mikolov 13] によって提案された，単語を多次元の分散表現ベクトルに落とし込む手法である．類似度については，コサイン類似度を使用して計算する．

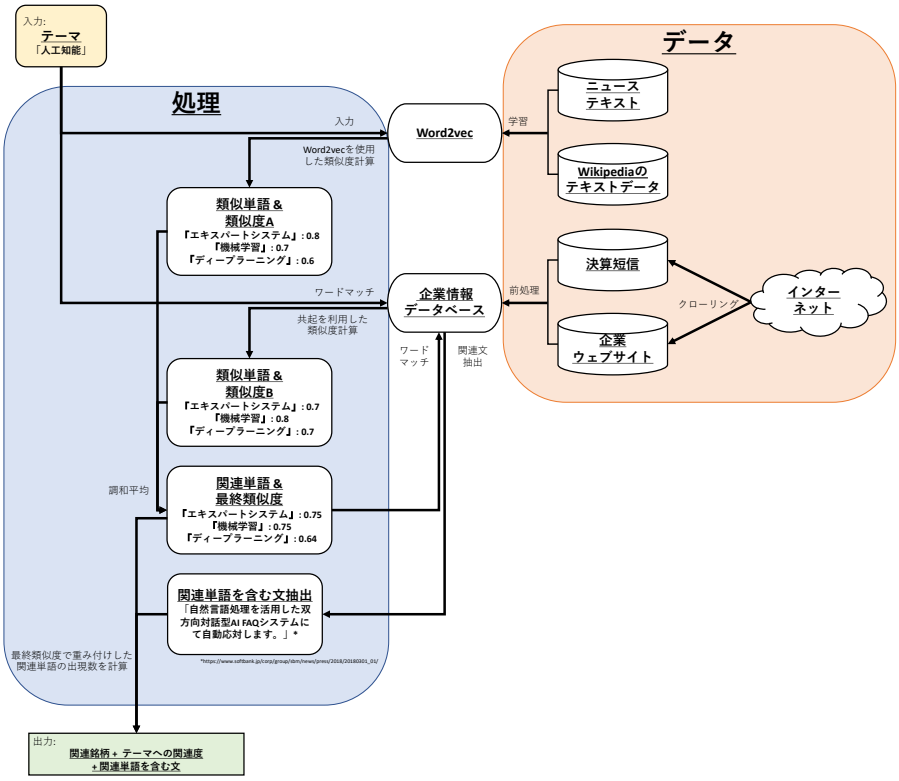


図 1: 手法概要図．例として,「人工知能」を入力とした時を載せている．ただし，数値などは実際のものと異なる．[Hirano 19a] より

Word2vec を利用した類似度計算において，異なるハイパーパラメーターセットで作成した Word2vec モデルを組み合わせて使用している．これは，Word2vec は学習が一意に同じ結果を返すわけではなく，かつ，ハイパーパラメーターによって結果に大きくばらつきが発生するため，それらの影響を緩和するために，異なるハイパーパラメーターセットを利用して学習させたモデルをアンサンブルした．主に，Dimensions，Word window size の異なるパラメーターセットで学習させた．

これらのハイパーパラメーターを使用して学習させたモデルを使用して，類似度を計算した．以下，この類似度を「類似度 A」と呼ぶ．類似度 A の計算方法は以下の通りである．

$$s'_{M_j,word_i} = \begin{cases} c_{M_j,word_i} & (\mathrm{word_i}\text{が}\ \mathrm{M_j}\text{における類似度上位 n 個に入る場合}) \\ 0 & (\text{他}) \end{cases} \quad (1)$$

ここで，$word_i$ は，全ての語彙の中で i 番目の単語，M_i は i 番目の Word2vec モデル，$c_{M_j,word_i}$ はテーマの単語として入力された単語と $word_i$ の M_j のモデルにおけるコサイン類似度，$s'_{M_j,word_i}$ は $word_i$ の M_j における類似度，$s_{A,word_i}$ は $word_i$ の類似度 A である．上記で定義された $s'_{M_j,word_i}$ の調和平均を取ることで，$s_{A,word_i}$ を計算する．さらに，$s'_{M_j,word_i}$ が 0 である場合は $s_{A,word_i}$ も 0 となる．つまり，Word2vec のモデルのいずれかにおいて，$word_i$ が類似度の上位 n 個に入っていない場合に $s_{A,word_i}$ も 0 となる．

この一連のアンサンブル的手法は，Nagata らの研究 [Nagata 18] を参考にしているが，新しく調和平均を取り入れ，拡張を行なった．

3.2 企業情報内における単語共起を利用した類似度計算

本手法において使用している二つの類似度のうち一つは 3.1 節で説明した通り，Word2vec を使用したものであり，もう一つの類似度を本節で説明する．

もう一つの類似度計算は企業情報内における単語共起を利用した類似度計算である．Word2vec は文脈に焦点を当てた手法であり，そのため,「東京」と「大阪」が類似単語として判定されるといった特徴がある．しかし，今回の目的に照らし合わせると，こういった，同じ文脈で使用されるが，関係性の薄い単語が含まれるなどといったことは望ましいとは言えない．そこで，企業情報内における単語共起を利用した類似度計算の手法を新たに提案するとともに，Word2vec と併用するこ

とにした．なお，この手法を利用することで，より下位の概念が強く取れることは，[Hirano 19b] で示している．

企業情報内における単語共起を利用した類似度計算により計算される類似度 B($s_{B,word_i}$) は次の通り計算される．図 2 は計算の具体例を示している．

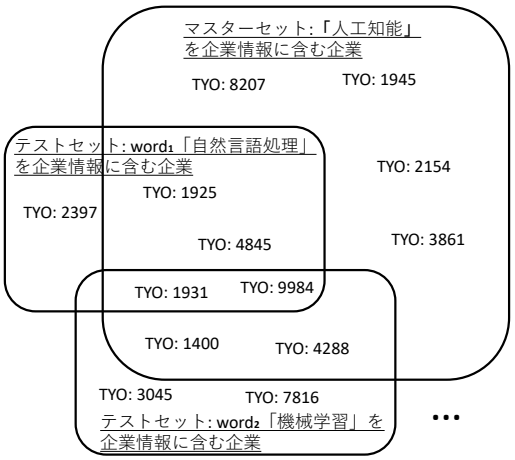


図 2: 類似度 B($s_{B,word_i}$) の計算例. “TYO: xxxx” は東京証券取引所におけるそれぞれの企業の銘柄コードを示している．ここで，$s_{B,word_1} = 4/5 = 0.8$，$s_{B,word_2} = 4/6 = 0.6667$ と計算される．これらの例はあくまで一部であり，実際の結果とは異なる．[Hirano 19a] より．

まず，インターネットを通じてクローリングにより取得した決算短信 (IR) と企業ウェブサイトから取得した企業情報を利用し，テーマの単語を企業情報に含む企業を抽出する．図 2 の例では，テーマの単語「人工知能」に対して，全部で 10 個の企業を抽出できている．(この例はあくまで一部であり，実際の結果とは異なる．) ここで，“TYO: xxxx” は東京証券取引所におけるそれぞれの企業の銘柄コードを示している．以下，このテーマの単語を元に抽出された企業のリストを「マスターセット」と呼ぶ．次に，ほぼ同等の操作を $word_1, word_2, \cdots$ で実施する．違いはターゲットとする単語がテーマの単語ではなく，$word_i$ になることである．つまり，$word_i$ を企業情報に含む企業を抽出していくことになる．図 2 の例では，$word_1$「自然言語処理」を含む企業は 5 つである．(この例もあくまで一部であり，実際の結果とは異なる．) ここで，これらの企業のリストを「テストセット ($word_1$)」と呼ぶことにする．同様にテストセット ($word_i$) を全て計算する．そして，最終的に $s_{B,word_i}$ を次式に従って計算する．

$$s_{B,word_i} = \frac{|\{\text{ マスターセット }\} \cap \{\text{ テストセット } (word_i)\}|}{|\{\text{ テストセット } (word_i)\}|} \quad (2)$$

これは，テストセット ($word_i$) のマスターセットに対する再現率の計算である．例えば，$word_1$「自然言語処理」の場合，5 つの企業がテストセットに含まれ，そのうち，4 つの企業だけがマスターセットにも含まれている．そのため，$s_{B,word_1}$ は $4/5 = 0.8$ と計算される．他も同様に計算されるので，$s_{B,word_2} = 4/6 = 0.6667$ となる．(これらの例もあくまで例であり，実際の結果とは異なる．)

3.3 最終類似度の計算および最終的な関連単語の抽出

以上で説明した二つの類似度を使用して，最終類似度 FS_{word_i} の計算を行う．最終類似度 FS_{word_i} は単なる $s_{A,word_i}$ と $s_{B,word_i}$ の調和平均により計算される．

次に，最終類似度 FS_{word_i} を用いて，最終的な関連単語の抽出を行う．最終類似度 FS_{word_i} の上位 n 個のみを最終的な関連単語とする．以下の説明においては，$word_{f_1}, word_{f_2}, \cdots, word_{f_k}$ を抽出された最終的な関連単語とする．

3.4 関連銘柄の抽出およびその関連度の計算

$word_{f_1}, word_{f_2}, \cdots, word_{f_k}$ に加え，$word_{f_0}$ をテーマの単語とし，$FS_{word_{f_0}}$ を 1 とする．それぞれの企業の企業情報の中から，$word_{f_0}, word_{f_1}, \cdots, word_{f_k}$ それぞれが含まれる文章を抽出し，それぞれの単語の出現回数をカウントする．そして，企業のテーマへの関連度 CS を次式の通り定義する．

$$CS = \sum_{i=0}^{k} FS_{word_{f_i}} \cdot count_{word_{f_i}} \quad (3)$$

ここで，$count_{word_{f_i}}$ は，それぞれの企業の企業情報の中での $word_{f_i}$ の出現回数である．この定義に基づき，各企業のテーマへの関連度が計算され，その関連度に基づいてランキングされる．

4 データと前処理

本章では，本研究において使用したデータと実施した前処理を説明する．本研究においては，日本語の文書を使用している．第一目標が日本市場をターゲットとしたテーマ型投資信託の作成をサポートすることであったからである．日本語の場合，英語などと異なり，形態素間にスペーシングが行われていないため，形態素解析を行う必要がある．

4.1 使用したデータ

本研究において，目的に応じていくつかのデータを使用した．目的は主に二つで，Word2vec の学習用および企業情報としてである．Word2vec 学習用として，ライブドアニュースコーパス[1]，Wikipedia 日本語記事 (version 21-Jun-2018 22:09) [2]，日本経済新聞 (1990–2015 and 2017; 2016 年分は技術的問題から使用していない) を利用した．これらのデータを元に，1,147,973 語彙，1,809,736,365 文字を含むテキストデータを Word2vec の学習に利用した．

企業情報としては，2012/12/9 — 2018/5/11 の決算短信，全 90,813 ファイル (PDF) を日本取引所グループ適時開示情報閲覧サービス (TDnet) [3] より取得した．企業ウェブサイトのデータとしては，2018/6/6 — 2018/6/25 の期間に，2,293,460 ファイルのみ (703,699 PDF，1,472,317 HTML，その他) を取得した．

これらの企業情報はインターネットを通じて取得した．本研究においては，これらの多種のデータを組み合わせることで，関連銘柄を抽出している．

4.2 実施した前処理

日本語の文書においては形態素解析が必要である，日本語の形態素解析器として，KyTea [Neubig 11] や JUMAN++ [Morita 15] などがあげられる．本研究においては，もう一種類の MeCab (version 0.996)[4]を使用した，また，MeCab 用の追加の日本語辞書として，NEologd[5]を使用した．

MeCab および NEologd を使用して，全てのテキストデータを形態素に分解してから使用している．Word2vec の学習においては，形態素解析して，形態素間にスペースを入れたデータを Word2vec の学習用テキストとして使用した．一方，企業情報は検索を早くするために，形態素解析後のテキストを，どの企業の情報なのかやデータのソースなどとともにリレーショナルなデータベースに保存して使用した．特に，収集した企業情報は 600GB 程度あり，実験において，リレーショナルな形で保存されていないと照会が非常に困難であるため，リレーショナルなデータベースでの保存を採用した．

5 実験と結果

本章においては，本手法のハイパーパラメーターチューニングと評価実験およびそれらの結果について述べる．

各節ごとでの説明の前に，実験で共通で使用している評価用データについて述べる．評価用データとは，提案手法等を使用して出てきた結果を評価するために使ったデータである．評価用に使用したテーマは「美容」,「育児」,「ロボット」,「娯楽」の 4 つである．経験のあるファンドマネージャー 4 名のタグ付けデータを使用し，それを元に評価用データを作成した．TOPIX500 の中からランダムに選ばれた 100 銘柄を使用し，タグ付けを行なってもらった．タグ付けの基準は同様で，以下の通りである．

0. 全く関係ない
1. あえて言えば関連している
2. この企業の事業の一部に関係がある
3. 非常に関係があり，この企業の代表的な事業だ

ただし，一つだけ異なるのは，ファンドマネージャーは自分のタグ付けに自信がない場合には，その旨のタグ付けを行うことができる．

表 1: ファンドマネージャーのタグ付けの例．“+” は銘柄のタグ付けに自信がないことを示す．

銘柄コード	FM1 タグ	FM1 自信なし	FM2 分類	FM2 自信なし
4544	1	+	0	
4555	1	+	0	
4578	1	+	0	
4661	2		0	
4665	1		3	
4676	1		0	

表 1 はファンドマネージャーによるタグ付けの例である．例えば，銘柄コード 4544 のみらかホールディングスに対して，ファンドマネージャー 1 は自信がないものの,「あえて言えば関連している」とタグをつけた．一方で，ファンドマネージャー 2 は「全く関係ない」とタグをつけた．このファンドマネージャー 4 名によるタグ付けに基づいて，評価用データを作成する．関連銘柄の判定は非常に難しいタスクである．(詳細については [Hirano 19c] を参照．) そのため，評価用データにおいては，1 名以上のファンドマネージャーが関連のある銘柄だと判断した場合には，関連銘柄とすることとした．これは，人間が行う以上，知識不足や見落としにより，関連のない銘柄だと判定してしまう可能性が充分に存在するからである．

そこで，ファンドマネージャーによるタグ付けを評価用データに変換する際の最終的な基準を次のように定めた．

[1] https://www.rondhuit.com/download.html#ldcc
[2] https://dumps.wikimedia.org/jawiki/latest/
[3] https://www.jpx.co.jp/equities/listing/tdnet/
[4] http://taku910.github.io/mecab/
[5] https://github.com/neologd/mecab-ipadic-neologd

表 2: ファンドマネージャーによるタグ付けの評価用データへの変換例．“FM” はファンドマネージャーを，“+” は自信がないことを示す．

	FM によるタグ付け				評価用データ	
銘柄	FM1	FM2	FM3	FM4	関連銘柄	関連度レーティング
A	1	3	2	1	✓	$(1+3+2+1)/4 = 1.7500$
B	0	1	0	0	✓	$(0+1+0+0)/4 = 0.2500$
C	0	1(+)	0	0	-	$(0+1\times0.5+0+0)/3.5 = 0.1429$
D	2	1(+)	0(+)	0	✓	$(2+1\times0.5+0\times0.5+0)/3 = 0.8333$
E	0	0	1(+)	0(+)	-	$(0+0+1\times0.5+0\times0.5)/3 = 0.1667$
F	0	0	1(+)	1(+)	-	$(0+0+1\times0.5+1\times0.5)/3 = 0.3333$

- 自信のないファンドマネージャーを除き，1 名以上のファンドマネージャーが関連銘柄だと判定した (1,2,3 のいずれかにタグ付けした) 場合は関連銘柄としてみなす．
- 4 名のファンドマネージャーの 0-4 の関連度のタグ付けを算術平均したものをその銘柄の関連度レーティングとする．ただし，自信のないファンドマネージャーが存在する場合は，平均を取る際にその人の重み付けを 1 ではなく，0.5 とする．

表 2 はファンドマネージャーによるタグ付けから評価用データへの変換例である．銘柄 A と B は典型的な例である．全てのファンドマネージャーが銘柄 A は関連銘柄とみなしている．そのため，評価用データにおいても，関連銘柄としている．銘柄 B も評価用データにおいては関連銘柄となっている．これは，1 名のファンドマネージャーが関連銘柄だとみなしているからであり，すでに述べたとおり，1 名ファンドマネージャーのみが関連銘柄とみなしている場合でも，他のファンドマネージャーが知識不足や見落としなどから関連銘柄としていない可能性が充分にあると考えられるからである．一方で，銘柄 C は少し異なる．1 名のファンドマネージャーが関連銘柄としているが，そのファンドマネージャーは自信がないとしている．このような場合には，評価用データにおいては，関連銘柄と判定しない．さらに，銘柄 D-F はもっと特殊なケースである．銘柄 C においても同様ではあるが，自信のないファンドマネージャーは 0.5 名分としてカウントして平均をとったものを関連度レーティングとしている．さらに，銘柄 E と F においては，関連銘柄とタグをつけているファンドマネージャーが全員自信がない場合のため，評価用データにおいては関連銘柄とは判定していない．

このように基準を定めたものの，自信がないとタグをつけられたデータは $28/1200 = 2.3\%$ しかなく，実際には 2 名以上が自信がないことを示すタグを同一テーマ，同一銘柄につけた銘柄 D-F のようなケースは存在しなかった．

5.1 ハイパーパラメーターチューニング

より良いハイパーパラメーターを探すために，ハイパーパラメーターチューニングを行なった．企業情報のデータソース，3.1 節で説明した複数の Word2vec のモデルをアンサンブルするフェーズにおける上位何個を取るかという値 (Hyperparameter1 という)，3.3 節で説明した，最終類似度の計算における上位何個を取るかという値 (Hyperparameter2 という) を変更して，実験をし，その結果からチューニングを行なった．ハイパーパラメーターのチューニングの方式はグリッドサーチであり，それぞれのパラメーターに対していくつかの離散的な値を使用して，その全ての組み合わせを実験し，その中で良い結果を出したパラメーターセットを使用するという手法である．グリッドサーチで使用したそれぞれのパラメータに対するグリッドは以下の通りである．

- 企業情報のデータソース: (1) 決算短信のみを使用，(2) 企業ウェブサイトからのデータのみを使用，(3) 決算短信および企業ウェブサイトからのデータの両方を使用
- Hyperparameter1: top-10, top-20, top-50, top-100, top-200, top-500, top-1000, top-2000
- Hyperparameter2: top-5, top-10, top-20, top-50, top-100, top-200, top-500, top-1000

評価用データとしては,「美容」,「育児」,「ロボット」,「娯楽」があるが，そのうち,「美容」,「育児」,「娯楽」の三つをハイパーパラメーターのチューニングに利用し，残りの一つの「ロボット」を性能評価用に使用した.「ロボット」を性能評価用に使用した理由は，事前の実験で，最も低い結果を出していたからである．

ハイパーパラメーターのチューニングにあたっては，結果の良し悪しをいちいち人間が判断することは不可能であり，何かしらの「良さ」の指標を利用しなければならない．そこで，本実験においては，F1 による

チューニングを採用することとした．結果として，採用したハイパーパラメーターセットは企業情報のデータソースが決算短信のみ，hyperparameter1 が top-500，hyperparameter2 が top-50 となった．

5.2 テストデータでの性能評価

前節で定めたハイパーパラメーターを採用し，残っている一つの「ロボット」の評価用データを使用して，性能評価を行なった．比較として，企業情報として決算短信のみを使用し,「ロボット」という単語をそれぞれの企業情報に単純照合をすることで企業を抽出するという実験を行なった．結果は表 3，4，5 の通りである．

表 3: 「ロボット」の評価用データとチューニング後のハイパーパラメーターを使用した場合の提案手法の結果に対する混同行列

		評価用データにおいて 関連銘柄	非関連銘柄
実験結果	関連銘柄	48	34
	非関連銘柄	9	9

表 4: 企業情報として決算短信のみを使用して,「ロボット」という単語をそれぞれの企業情報に単純照合をすることで企業を抽出した場合の結果に対する混同行列 (比較用)

		評価用データにおいて 関連銘柄	非関連銘柄
実験結果	関連銘柄	9	5
	非関連銘柄	48	38

表 5: テスト用テーマ「ロボット」における，提案手法と比較手法の Precision，Recall，F1，Accuracy．比較手法は企業情報として決算短信のみを使用して,「ロボット」という単語をそれぞれの企業情報に単純照合をすることで企業を抽出するという手法である．

	Precision	**Recall**	**F1**	**Acc.**
提案手法	0.5854	0.8421	0.6906	0.5700
比較手法	0.6429	0.1579	0.2535	0.4700

表 5 を見れば明らかであるが，表 3 と 4 を比べると，比較手法に比べて，より多くの銘柄を関連銘柄として抽出しているため，Precision が低下している．一方で，圧倒的に Recall が向上しており，その結果，F1 も向上している．また，単純な結果の正確さを測る Accuracy も向上しており，提案手法が良い結果を出していることがわかる．

6 考察

本研究は，テーマ関連銘柄の抽出を支援するシステムの構築を目標にしたものであった．[Hirano 19c] で示しているが，これは，経験のあるファンドマネージャーにとっても難しいほどのタスクであった．そのため，そもそも評価用データがどれだけ正しいのか，という点については多少の疑問が残る．しかし，表 5 を見る限りでは，基本的な手法を利用した比較手法よりは良い結果を出しており，提案手法の一定の有効性が示されたと考えられる．

今回の結果は企業ウェブサイトからのデータはなくても，決算短信だけでも充分であるという結果になっている．しかしながら，企業ウェブサイトのデータを使用した場合には，一つの単語で抽出できる企業数が多いこともわかっている．これは，決算短信だけでなく，企業ウェブサイトからのデータも使った場合には，少ない数の関連単語のみを使用することで，充分な関連銘柄を抽出していた一方で，決算短信だけを企業情報で使用する場合には，より多くの関連単語を採用することで，データの少なさをカバーすることができているのかもしれない．しかし，実験で実際に使用した 4 つの単語はすでに広く普及している，一般的な単語であった．そういった一般的単語は決算短信には頻出する一方で，比較的新しい単語は決算短信には出現しにくい可能性がある．そのため，今後の課題として，提案手法において，企業ウェブサイトからのデータを使用しなかった場合に，比較的新しい単語においても充分な精度を発揮するのかを確認しなければならない．

今回の実験における，ハイパーパラメーターのチューニングはある種の教師あり学習のような形のグリッドサーチで行なった．教師データとしては，ファンドマネージャーにタグ付けしてもらったデータを元に作成した評価用データを利用したが，このデータがどれだけ正しいかわからないだけでなく，作成も人手で行うため，時間のかかる作業であり，多くの評価用データを作成するのは困難であった．そのため，評価に使用したデータのテーマは限られており，これでは充分な結果とは言えない可能性もある．もちろん，より正確で，多くの評価用データを作成することができれば問題はないが，限界がある．さらに，現在は日本株だけを対象にしていたが，これを世界中の株を対象にすると，評価用データの作成の難易度はもっと上がると考えられる．そのため，教師なし学習または半教師あり学習のような手法を作成しなければ，どんなテーマに

も対応できるような手法とはならないと考えられる．

今後の展望としてだが，現時点では，単純にテーマへの関連度のみを考慮してランキングを作成していたが，実際にはもう少し他の軸を加えてランキング付けしたいという需要もある．例えば，テーマにおける代表性や企業の大きさを加味したランキングが欲しいという意見もファンドマネージャーから聞いた．そのため，様々な指標を取り入れられるようなスキームを作成することも必要かもしれない．また，銘柄をテーマ型投資信託に採用し，運用を行う際には，ポートフォリオ運用をすることになるので，同じような株価変動をする銘柄ばかりを組み込むことは望ましくない．それは，単に価格変動だけではなく，似たような事業を行なっている企業を同時に取り込みことは，結果として，同じような価格変動を引き起こす可能性を秘めている．そのため，価格変動や企業情報を主成分分析したり，ベクトル化することで，様々な需要にうまく適用できるようにする技術の開発などが必要であると考える．

謝辞

本研究は，大和証券グループと東京大学大学院工学系研究科により開設された社会連携講座「次世代運用テクノロジー」における，大和証券投資信託委託株式会社および株式会社大和総研との共同研究の一部であり，大和証券グループの支援を受け，行われました．大和証券グループ各社の多くの方にご支援をいただきました．この場を借りて御礼申し上げます．

参考文献

[Fellbaum 98] Fellbaum, C.: *WordNet: An Electronic Lexical Database*, The MIT Press (1998)

[Hirano 18] Hirano, M., Sakaji, H., Kimura, S., Izumi, K., Matsushima, H., Nagao, S., and Kato, A.: Selection of Related Stocks using Financial Text Mining, in *Proceedings of 18th IEEE International Conference on Data Mining Workshops (ICDMW)*, pp. 191–198, Singapore, Singapore (2018)

[Hirano 19a] Hirano, M.: *Extraction of Related Stocks for Themed Mutual Funds using Text Mining*, Graduation thesis, The University of Tokoy (2019)

[Hirano 19b] Hirano, M., Sakaji, H., Kimura, S., Izumi, K., Matsushima, H., Nagao, S., and Kato, A.: 文書内における単語の共起を利用した上位下位概念の推定, 言語処理学会第 25 回年次大会, pp. 597–600, Nagoya, Aichi, Japan (2019), The Association for Natural Language Processing

[Hirano 19c] Hirano, M., Sakaji, H., Kimura, S., Izumi, K., Matsushima, H., Nagao, S., Kato, A., Hirano, M., Sakaji, H., Kimura, S., Izumi, K., Matsushima, H., Nagao, S., and Kato, A.: Related Stocks Selection with Data Collaboration Using Text Mining, *Information*, Vol. 10, No. 3 (2019)

[Ito 18a] Ito, T., Sakaji, H., Izumi, K., Tsubouchi, K., and Yamashita, T.: GINN: gradient interpretable neural networks for visualizing financial texts, *International Journal of Data Science and Analytics*, pp. 1–15 (2018)

[Ito 18b] Ito, T., Sakaji, H., Tsubouchi, K., Izumi, K., and Yamashita, T.: Text-Visualizing Neural Network Model : Understanding Online Financial, in *Proceedings of the 22nd Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2018)*, pp. 247–259, Melbourne, Australia (2018)

[Kitamori 17] Kitamori, S., Sakai, H., and Sakaji, H.: Extraction of sentences concerning business performance forecast and economic forecast from summaries of financial statements by deep learning, in *Proceedings of 2017 IEEE Symposium Series on Computational Intelligence (IEEE SSCI 2017)*, pp. 67–73, Honolulu, Hawaii, USA (2017)

[Koppel 06] Koppel, M. and Shtrimberg, I.: Good News or Bad News? Let the Market Decide, in *Computing Attitude and Affect in Text: Theory and Applications*, pp. 297–301, Springer Netherlands (2006)

[Low 01] Low, B.-T., Chan, K., Choi, L.-L., Chin, M.-Y., and Lay, S.-L.: Semantic expectation-based causation knowledge extraction: A study on Hong Kong stock movement analysis, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2001)*, pp. 114–123 (2001)

[Mikolov 13] Mikolov, T., Chen, K., Corrado, G., and Dean, J.: Efficient Estimation of Word Representations in Vector Space, in *Proceedings of the International Conference on Learning Representations (ICLR 2013)*, pp. 1–12, Scottsdale, Arizona, USA (2013)

[Milea 10] Milea, V., Sharef, N. M., Almeida, R. J., Kaymak, U., and Frasincar, F.: Prediction of the MSCI EURO index based on fuzzy grammar fragments extracted from European Central Bank statements, in *2010 International Conference of Soft Computing and Pattern Recognition*, pp. 231–236 (2010)

[Morita 15] Morita, H., Kawahara, D., and Kurohashi, S.: Morphological Analysis for Unsegmented Languages using Recurrent Neural Network Language Model, in *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP 2015)*, pp. 2292–2297, Lisbon, Portugal (2015)

[Nagata 18] Nagata, R., Nishite, S., and Ototake, H.: A Method for Detecting Overgeneralized Be-Verb based on Subject-compliment Identification [published in Japanese], *The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018 (JSAI 2018)* (2018)

[Neubig 11] Neubig, G., Nakata, Y., and Mori, S.: Pointwise Prediction for Robust , Adaptable Japanese Morphological Analysis, in *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL HLT 2011)*, pp. 529–533, Portland, Oregon, USA (2011)

[Sakai 07] Sakai, H. and Masuyama, S.: Extraction of Cause Information from Newspaper Articles Concerning Business Performance, in *Proceedings of the 4th IFIP Conference on Artificial Intelligence Applications & Innovations (AIAI 2007)*, pp. 205–212 (2007)

[Sakaji 08] Sakaji, H., Sakai, H., and Masuyama, S.: Automatic Extraction of Basis Expressions That Indicate Economic Trends, in *Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD 2008)*, pp. 977–984 (2008)

[Sakaji 17] Sakaji, H., Murono, R., Sakai, H., Bennett, J., and Izumi, K.: Discovery of Rare Causal Knowledge from Financial Statement Summaries, in *The 2017 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr 2017)*, pp. 602–608 (2017)

[Schumaker 09] Schumaker, R. P. and Chen, H.: Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFin Text System, *ACM Transactions on Information Systems*, Vol. 27, No. 2, pp. 12:1–12:19 (2009)

[Xing 18] Xing, F., Cambria, E., and Welsch, R. E.: Growing Semantic Vines for Robust Asset Allocation (2018), https://ssrn.com/abstract=3275132

文脈を考慮した決算短信からの業績要因抽出

Extraction of corporate performance factors from summaries of financial statements considering context

加藤悠太 [1]　　酒井浩之 [1*]
Yuta Kato[1]　　Hiroyuki Sakai[1]

[1] 成蹊大学 理工学研究科 理工学専攻
[1] Department of Computer and Information Science, Faculty of Science and Technology, Seikei University

Abstract: In this paper, we proposed a method for automatically extracting sentences containing corporate performance factors from summaries of financial statements at high accuracy. More specifically, only sentences that can be determined to be corporate performance factors in the summaries of financial statements with high probability are extracted, and those sentences are used as training data. We trained the neural network by using the word representation of the training data. Furthermore, by using the appearance tendency of the corporate performance factor sentence specific to summaries of financial statements as a bias of model output, it became possible to extract with a higher f-measure than related work that performs filtering processing using corporate keywords .

1　はじめに

近年，証券市場における個人投資家の比重が増加しており個人投資家への投資判断を支援する技術の必要性が高まっている．そのため，人工知能分野における手法を金融市場の様々な場面に応用することが期待されている．例として，日本銀行が毎月発行している「金融経済月報」や経済新聞記事を解析し，経済市場を分析する研究などが盛んに行われている [1][2].

投資家にとって，企業の業績情報を収集することは重要であり，その中でも業績の要因となる情報が重要である [3][4]．しかしながら，証券市場における上場企業は膨大な数存在し，その企業が四半期に 1 回，決算発表を行う．そのため，人手により多くの企業の業績要因を把握することは困難である．決算短信の文において，この業績の要因が含まれている文を，本研究では「業績要因文」と定義する．

決算短信から業績要因文を抽出する手法とし，既存手法では特定の文を見つける上で手がかりとなる表現である「手がかり表現」を拡張した「拡張手がかり表現」と，企業の重要なキーワード「企業キーワード」による重みを用いて業績要因文を抽出する．その手法は適合率のみは非常に高い結果で抽出することが可能なため，その抽出した業績要因文を学習データの正例として深層学習を行う．この学習モデルによる業績要因文の抽出を行い，その抽出された文に対し企業キーワードによるフィルタリングをする手法が提案されている [5].

しかしながら，既提案手法は深層学習による文の抽出後，企業キーワードによるフィルタリング処理を行うことによって適合率を向上させているため，新規事業のような過去の決算短信で出現していなかったキーワードが含まれる業績要因文を抽出することができない．

そのため，本手法では企業の重要なキーワードによるフィルタリング処理を行わず，業績要因文の文脈や前後関係を考慮した手法によって既提案手法と同程度の適合率，再現率で業績要因文の抽出をすることを目的とする．

2　関連研究

既存手法 [6] を用い，決算短信から業績要因文，手がかり表現，企業キーワードを抽出する．そしてさらに手法 [5] により手がかり表現から拡張手がかり表現を獲得し，抽出された業績要因文に対し企業キーワードを用いてスコアを付与する．そのスコアと拡張手がかり表現により学習データの自動生成を行う．生成された学習データを用い，深層学習（多層パーセプトロン）

*連絡先：成蹊大学
〒 180-8633 東京都武蔵野市吉祥寺北町 3-3-1
E-mail: h-sakai@st.seikei.ac.jp

にて決算短信から業績要因文を抽出する．既存手法の概要を図 1 に示す．

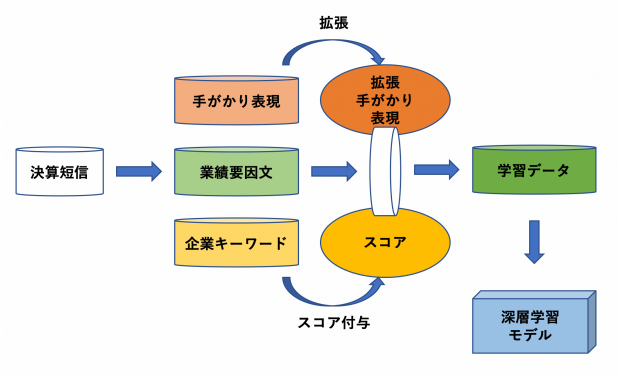


図 1: 既存手法 [5] の概要図

2.1 学習データの自動生成

獲得された拡張手がかり表現と企業キーワードを用い，以下の手順により学習データの自動生成を行う．

Step1: 決算短信から手法 [6] により業績要因文を抽出する．

Step2: 抽出した業績要因文に対し，企業キーワードを含みかつ拡張手がかり表現を含む文を抽出する．

Step3: 業績要因文に含まれている企業キーワードのスコアの合計をその業績抽出文のスコアとして付与する．

Step4: 業績要因文に付与したスコアの企業ごとの平均値を算出し，平均値より大きいスコアをもつ業績要因文を学習データの正例データとする．

Step5: 企業の決算短信から企業キーワードと手がかり表現のどちらも含まない文を抽出する．

Step6: Step5 で抽出した文のうち文字数が一定以上の文を学習データの負例データとする．

2.2 特徴量選択

自動生成した学習データを用い，業績要因文を抽出する．学習データにおいて正例の業績要因文に含まれる内容語（名詞，動詞，形容詞）に対し，以下の式にて重みを計算する．

$$W_p(t, S_p) = TF(t, S_p)H(t, S_p) \quad (1)$$

S_p: 正例に属する業績要因文の集合．

$TF(t, S_p)$: 文集合 S_p において語 t が出現する頻度．

$H(t, S_p)$: 文集合 S_p における各業績要因文に含まれる語 t の出現確率に基づくエントロピー．

$$H(t, S_p) = -\sum_{s \in S_p} P(t, S_p) \log_2 P(t, S_p) \quad (2)$$

$$P(t, S_p) = \frac{tf(t, s)}{\sum_{s \in S_p} tf(t, s)} \quad (3)$$

ここで，$P(t, s)$ は業績要因文 s における語 t の出現確率を表し，$tf(t, s)$ は文 s において語 t が出現する頻度を表す．

次に，負例の文に含まれる内容後（名詞，動詞，形容詞）に対しても同様に重みを計算する．

$$W_n(t, S_n) = TF(t, S_n)H(t, S_n) \quad (4)$$

ここで，S_n は学習データにおいて負例に属する文の集合を表す．

以下の条件どちらかが成り立つ語 t を特徴量として選択する．

$$W_p(t, S_p) > 2W_n(t, S_n) or W_n(t, S_n) > 2W_p(t, S_p) \quad (5)$$

2.3 業績要因文の抽出

多層パーセプトロンで学習されたモデルにより決算短信から業績要因文を抽出する．その後，精度を高めるために，抽出された文に含まれている企業キーワードのスコアの合計をその文のスコアとして付与し，スコアが高いかつ文末が「た」文のみを業績要因文として抽出する..

3 分散表現

3.1 単語の分散表現

既存手法 [5] では単語の重みを計算し特徴量選択を行っていたが，本手法では新規事業のような未知の単語が出現する文や業績要因文の中でも判定が曖昧な文に対し分類を行うことを可能にするため，単語の分散表現を用いることにより精度向上を図る．具体的には，分散表現獲得の基礎的な手法として，word2vec[7][8] と，より単語の文脈的な解釈を表すことが可能な ELMo[9] を使用した． 本手法では，word2vec の学習は Skip-gram により行った．

ELMo のモデル構造は双方向の LSTM から構成されており，単語の文脈情報を獲得することができる．例えば「増加」といった単語のベクトルを得たい場合，word2vec では「増加」という単語の周辺単語の出現確率から学習を行い，その結果「増加」という単語の分

散表現を一意で持つことになる．しかし ELMo では単語のコンテキストを考慮することにより，その単語が文脈的にどう扱われるのかという情報を分散表現として表すことが可能となる．そのため,「増加」という単語の文中での意味を表現することが可能となる．

3.2 文の分散表現

単語以外では文に対する分散表現の研究もされている．比較対象としてトピックを考慮した SCDV[10] を用いた．SCDV は単語がトピッククラスタに属する確率を GMM による分類から算出し，その確率と単語の *idf* 値から文の分散表現を生成し，その要素をスパースする手法である．

4 ELMo を用いた LSTM による業績要因文の抽出

ELMo は双方向 LSTM によって単語のコンテキストを学習した分散表現を獲得することが可能であり，それを word2vec のようなベクトルに連結することによって精度を上げることが可能とされている．また，一般的には ELMo 自体がコンテキストを学習しているため，獲得した分散表現に対して時系列性を考慮したモデルによって学習する必要はないとされている．しかしながら ELMo の学習は非常にコストが高いため，短期間でモデルの学習をすることは難しい．そのため，予め wikipedia の一般的なコーパスによって学習されたモデル [11][12] を用い，そのモデルの出力する分散表現を一般的な日本語におけるコンテキストと考え，さらにそのモデルの出力である決算短信の分散表現を LSTM によって学習を行う．これにより ELMo により決算短信を学習することに近づくと考えた．

本来 ELMo は双方向 LSTM によって学習を行うが，ELMo における決算短信の分散表現をさらに学習する際には，ある程度日本語としてのコンテキストは学習されていると判断し，図 2 のような順方向 LSTM のみで学習を行い，その出力を多層パーセプトロン（MLP）の入力として用いる．

5 業績要因文の出現傾向におけるバイアス値

決算短信の文集合において，業績要因が出現する頻度や傾向を分析した．その結果，図 3 のように業績要因文は連続して出現することが多いと分かった．この傾向は一度業績の要因について言及し始めると，それ

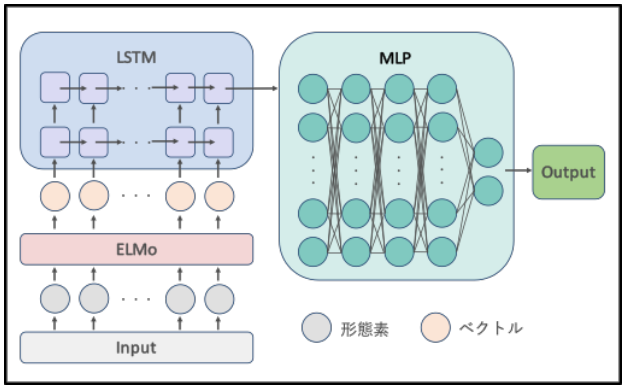


図 2: ELMo の分散表現を入力としたモデルの概要図

に関連する業績の要因をまとまった形で記すからであると考えられる．これを利用し業績要因が含まれてい

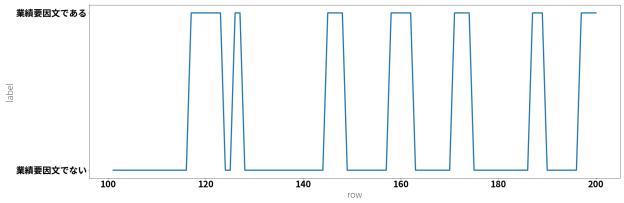


図 3: 業績要因文の出現傾向

ると判断された文の前後には業績要因が含まれている可能性が高いという仮説のもと，ある文の判定に対しその前後のモデルの出力をバイアスとしてかけることにより，決算短信における業績要因文の出現傾向の特徴を考慮した判定結果を得ることとした．これにより，決算短信に含まれる業績要因文の位置関係に基づく特徴をバイアス値として表現することができ，より決算短信に特化した分類結果を得ることが可能になる．このため，決算短信から業績要因文の抽出における精度向上が期待できる．

連続した決算短信の文 $s_1, s_2, \cdots, s_i, s_I$ があったとき，以下の式にて前後の文のバイアスをかけた判定結果 o'_i を得る．

$$o'_i = \frac{o_i + b_i}{2} \tag{6}$$

$$b_i = \frac{\sum_{-c \le j \le c, j \neq 0} 2^{|c|-|j|} o_{i+j}}{\sum_{-c \le j \le c, j \neq 0} 2^{|c|-|j|}} \tag{7}$$

このとき，

o_i: 文 s_i をモデルの入力とした時の出力結果．

b_i: 文 s_i に対し，前後の出力結果を加重平均したバイアス値．

c: コンテキストウィンドウ．決算短信の前後何文までをバイアスとして含めるか表すパラメータ．

これにより，業績要因文が集合している場合には業績要因文である可能性が高いというバイアスがかかり，また，業績要因が含まれていない文が集合している場合には業績要因が含まれていない可能性が高いというバイアスがかかることになり，決算短信における業績要因文の出現傾向の特徴を考慮した判定が可能となる．

6 実験

既存手法 [5] と本手法で結果を比較するため，自動生成された学習データを用い，word2vec，ELMo による単語の分散表現を特徴量とした MLP，LSTM[13][14] による業績要因文の抽出を行った．また，文の分散表現として SCDV を用いた MLP で業績要因文の抽出を行った．

word2vec，SCDV による分散表現を用いたモデルではハイパーパラメータ探索として k-fold 法を用いた 3 分割交差検証を行い，ベイズ最適化 [15] により探索を行った．ELMo に関しては計算コストが高いため，少量の汎用的なパラメータによってグリッドサーチを行った．

7 評価

7.1 分散表現を用いた業績要因抽出の評価

分散表現を用いた場合の業績要因の抽出結果について評価を行った．その評価結果を表 1 に示す．既存手法と比較し，ELMo を特徴量として LSTM と MLP により判定を行なった結果は，企業キーワードによるフィルタリング処理を行わなくとも，既存手法と同等の F 値を得ることができた．また，word2vec を特徴量として LSTM と MLP により判定を行なった結果に関しても既存手法には劣るが比較的良好な結果が得られた．

7.2 業績要因文の出現傾向をバイアスとした業績要因文の抽出

分散表現を用いた場合の業績要因抽出のモデルの出力に対し，決算短信の業績要因文出現傾向をバイアスとした実験について評価を行った．その評価結果を表 2 に示す．ELMo の分散表現を用いたモデルの出力にバイアスをかけた結果，コンテキストウィンドウ c を 1，2 に設定した場合，適合率，再現率共に向上する結果となった．

この結果から判定する前後の文との関係を考慮し，モデルの出力に対しバイアスとしてかけることにより精度が向上することがわかった．

表 1: 分散表現を用いた業績要因抽出の評価結果

	適合率	再現率	F 値
既存手法			
MLP			
企業キーワードフィルタ	0.827	0.724	0.772
特徴量: SCDV			
モデル: MLP	0.358	0.724	0.523
特徴量: word2vec			
モデル: MLP	0.462	0.859	0.601
特徴量: word2vec			
モデル: LSTM,MLP	0.609	0.834	0.704
特徴量: ELMo			
モデル: MLP	0.376	0.945	0.538
特徴量: ELMo			
モデル: LSTM,MLP	0.723	0.847	0.780

表 2: 出現傾向をバイアスとした抽出結果

	c	適合率	再現率	F 値
特徴量: ELMo モデル: MLP	0	0.376	0.945	0.538
	1	0.406	0.945	0.568
	2	0.413	0.949	0.575
特徴量: ELMo モデル: LSTM,MLP	0	0.723	0.847	0.780
	1	0.787	0.859	0.821
	2	0.796	0.876	0.835

8 考察

実験結果から，単語の文脈を考慮した分散表現を用いることにより，既存手法を上回る F 値を得ることができた．これは決算短信の 1 文における語彙数の分散が大きく，文の長さが 2 倍以上になる文も多く存在する．このため，tf 値やエントロピーなどの単語数による影響を受けやすい特徴量から分散表現にすることで，文の長さの影響を軽くすることが可能になったと考えられる．また，文脈による単語の扱われ方を考慮した表現に変えることで，業績の要因について書かれている文の単語のコンテキストを特徴として学習することが可能になったことも精度向上につながったと考える．

決算短信から業績要因文の抽出後，企業キーワードによるフィルタリング処理を行わなくすることで，学習データに存在しない企業キーワードになりうる単語が出現する文を決算短信から抽出することが可能になった．このため，企業が新規事業に対する業績要因として発表している文についてもフィルタリングによって除外されることはなくなった．また，新規事業の場合は新規で出現する単語が多く，重みの算出に影響がある．それを単語のコンテキストを考慮した分散表現に変えることや，新規事業に関する単語であっても，単

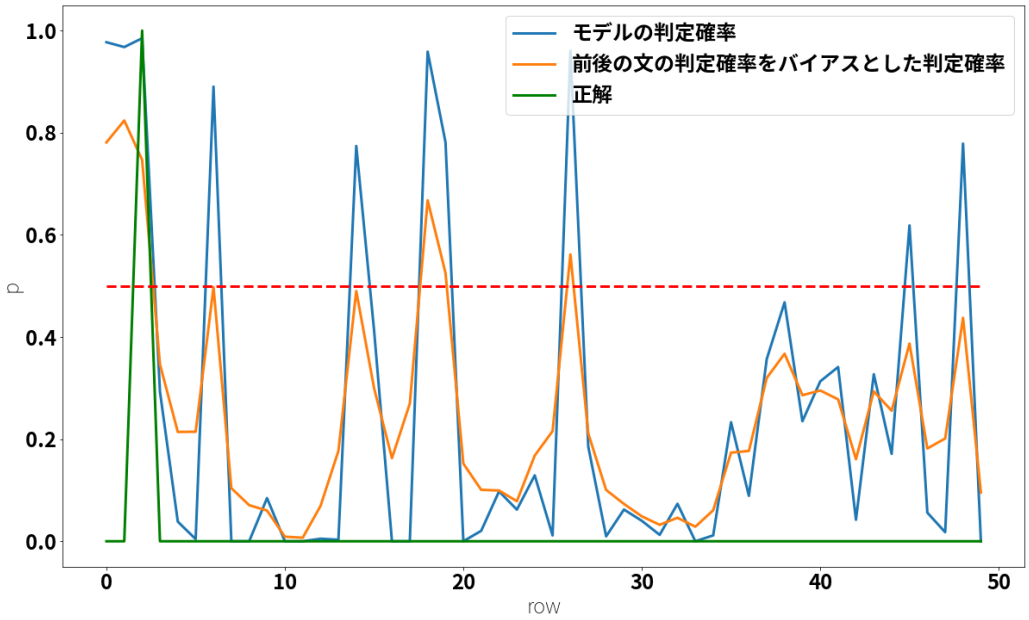


図 4: バイアスを用いた場合の判定結果の比較

語の分散表現を得るために使用した wikipedia には含まれていることは多いので，より新規事業に関する文を抽出することが可能になった一因と考えられる．

今回，文脈的解釈を含んだ分散表現が有用であると判断できたため，ELMo に SCDV のトピックのような観点を加えることでより精度が上がる可能性もある．そのため，トピックとコンテキストの双方を組み込むことが可能な分散表現についても決算短信の分析において有用であるか判断が必要である．

ELMo の分散表現については 1024 次元で行なったが，ベクトルの圧縮や学習データのサンプリングを行い，サンプリングしたデータで最適な学習パラメータを探索することも必要であると考えられる．

業績要因文であるかどうか判定する文の前後の判定結果を利用し，モデルの出力にバイアスをかける手法では有用な適合率，再現率が得られた．これは業績要因文であると誤判定してしまうことを前後の出力をバイアスとして用いた場合には抑制することが可能になったためである．図 4 からも前後の文の出力結果に偏り，誤判定が抑制していることが読み取れる．

前後の文が判定結果に依存するのであれば，単語や文字で使われる n-gram を文単位での文 n-gram として扱うことで判定結果が変わることになり，この結果を見ることで前後の文が判定文に対しどのように作用しているかを分析することが可能になると考えられる．近年では，bert[15] のような隣接文判定をタスクの一つとして解いているモデルもあるため，この検証も必要である．

9 むすび

本研究では，決算短信から業績要因を抽出する手法を企業の重要なキーワードや手がかり表現によるフィルタリングを行わず，文脈を考慮した手法によって既提案手法と同等の適合率，再現率で業績要因文の抽出を行った．具体的には，単語の頻度に基づく特徴量から単語の周辺確率やコンテキストを分散表現として表したものを利用し，ニューラルネットにより学習，判定を行うことで F 値 0.78 と既提案手法と同等の精度で抽出することが可能となった．さらに，決算短信における業績要因文の出現傾向の特徴として，業績要因が含まれる文は文章内で固まった状態で出現する性質を利用し，モデルの出力に対し，さらに前後のモデルの出力をバイアスとしてかけることで高い精度で抽出することができた．評価の結果，F 値が 0.83 と良好な結果を得ることができた．

今後の課題として，学習データと評価データの乖離を減らすことや，企業や決算短信ごとでの業績要因文の書き方を特徴として捉えることでバイアス値の調整や，学習するときの特徴量自体により前後の文の特徴を含ませることがあげられる．

参考文献

[1] 和泉潔, 後藤卓, 松井藤五郎: 経済テキスト情報を用いた長期的な市場動向の推定, 情報処理学会誌, Vol. 52, No. 12, pp. 3309–3315 (20011)

[2] 蔵本貴久, 和泉潔, 吉村忍, 石田智也, 中嶋啓浩, 松井藤五郎, 吉田稔, 中川裕志: 新聞記事のテキストマイニングによる長期市場動向の分析, 人工知能学会論文誌, Vol. 28, No. 3, pp. 291–296 (2013)

[3] Hiroyuki, Sakai., Shigeru, Masuyama.: Assigning Polarity to Causal Information in Financial Articles on Business Performance of Companies, *IEICE Transactions on Information and Systems*, Vol. E92-D, No. 12, pp. 2341–2350 (2009)

[4] Hiroyuki, Sakai., Shigeru, Masuyama.: Cause Information Extraction from Financial Articles Concerning Business Performance, *IEICE Transactions on Information and Systems*, Vol. ED, No. 4, pp. 959–968 (2008)

[5] 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)

[6] 酒井浩之, 西沢裕子, 松並祥吾, 坂地泰紀: 企業の決算短信 PDF からの業績要因の抽出, 人工知能学会論文誌, Vol. J98-D, No. 5, pp. 172–182 (2015)

[7] Tomas, Mikolov., Kai, Chen., Greg, Corrado., Jeffrey, Dean.: Efficient Estimation of Word Representations in Vector Space, *ICLR Workshop* (2013)

[8] Xin, Rong.: word2vec Parameter Learning Explained, *arXiv preprint arXiv:1411.2738*, (2014)

[9] Matthew, E, Peters †., Mark, Neumann., Mohit, Iyyer., Matt, Gardner., Christopher, Clark., Kenton, Lee., Luke, Zettlemoyer.: Deep contextualized word representations, *Proc. of NAACL*, (2018)

[10] Mekala, Dheeraj., Gupta, Vivek., Paranjape, Bhargavi., Karnick, Harish.: SCDV: Sparse Composite Document Vectors using soft clustering over distributional representations, *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pp. 659–669, (2017)

[11] Che, Wanxiang., Liu, Yijia., Wang, Yuxuan., Zheng, Bo., Liu, Ting.: Towards Better UD Parsing: Deep Contextualized Word Embeddings, Ensemble, and Treebank Concatenation, *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pp. 55–64, (2018)

[12] Fares, Murhaf., Kutuzov, Andrey., Oepen, Stephan., Velldal, Erik.: Word vectors, reuse, and replicability: Towards a community repository of large-text resources, *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pp. 271–276, (2017)

[13] S, Hochreiter., J, Schmidhuber.: Long short-term memory, *Neural Computation*, 1735-1780, (1997)

[14] Felix, A, Gers., Jurgen, Schmidhuber., Fred, Cummins.: Learning to forget: Continual prediction with LSTM *Neural Computation*, pp. 2451–2471, (2000)

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