日本株データベースにおける 曜日効果を考慮した投資モデルの網羅的分析

Exhaustive Analysis of Portfolios under the Day of the Week Effect in the Japanese Stock Market Database.

林 大祐 *[†] 羽室 行信 [‡] 岡田 克彦 [‡] 湊 真一 [†]
Daisuke HAYASHI Yukinobu HAMURO Katsuhiko OKADA Shin-ichi MINATO

北海道大学 情報科学研究科 [†]
Graduate School of Information Science and Technology, Hokkaido University

関西学院大学 経営戦略研究科 [‡]
Institute of Business and Accounting, Kwansei Gakuin University

概要

In the literature of financial economics, several calendar anomalies are reported including ”the day-of-the-week effect”. Especially in the weekend, traders tend to square out their positions on Friday, thus making the Monday returns susceptible to the news during the weekend. Prior studies find such day-of-the-week regularities based on the stock market averages. In this paper, we focus on the individual stocks' day-of-the-week skewness of returns and explore whether the alpha generating portfolio can be constructed exploiting such regularities.

1. 背景

近年ビックデータを金融分野に活用する技術が注目されており、株式にかかる様々なデータをマイニングする既存研究が現れている。大規模なテキスト解析する事で企業や投資家の行動心理を数値化する研究、また、投資家心理の揺れを利用しながらポートフォリオ構築する研究、銘柄別に旬の季節を探索しポートフォリオ構築する研究等である [1][2][3] 。本研究も大規模データを扱いながら金融の既存研究で明らかになっていない領域に光をあてる。具体的には、週単位のカレンダーアノマリーである「曜日効果」と呼ばれる現象に焦点をあて、超過リターンを獲得できるポートフォリオが構築可能かどうかを検証する。そもそも株式市場は企業業績を反映して価格形成がおこなわれているため、特定の曜日を取引の選択基準として採用することで超過収益を獲得できることはないはずである。しかし、株式市場全体として検証した場合、「曜日効果」が存在するという報告が世界の様々な株式市場についてなされており、金融分野のパズルとして扱われている [4][5]。ただ、平均値としての株価指数に現れる曜日効果は、個別銘柄としてみれば別の様相を示す可能性があり、個別銘柄に着眼した曜日効果の網羅的な研究は筆者らが知る限り存在しない。そこで、本研究では日本の全上場銘柄を対象に、個別銘柄それぞれについて超過リターンが得られるポートフォリオを構築できるかを網羅的に探索すると同時に、各ポートフォリオ毎にどの曜日効果が強いかを検証した。

* 連絡先: 林 大祐
北海道大学 情報科学研究科 大規模知識処理研究室
〒060-0814 北海道札幌市北区北 14 条西 9 丁目
Email: hayashi_da@ist.kokudai.ac.jp

2. 準備

2.1 株価データベース

本研究では、約 15 年間 (2001 年 1 月〜2015 年 5 月) の約 4000 銘柄についての株価データを使用。各銘柄別に時々刻々の取引データ (トランザクションデータ) ではなく、ここでは各銘柄の日々の株価データ（始値、最高値、最安値、終値）が csv 形式で保持されているデータベースを採用している。(総行数:1300 万行、総バイト数:6.6 億バイト)

2.2 ベンチマークと TOPIX 投資の収益率

新規の投資手法を提案する際、収益が多い事を示すだけでは投資家を納得させることは出来ない。市場全体の浮き沈み (景気の良し悪し) に左右されることなく収益が上がることを示す必要がある。ここで必要となるのが、特定のベンチマークである。ベンチマークとは、ある投資手法を評価するための指標である。TOPIX とは、時価総額¶により重み付けを行い、東証市場‖に上場している全銘柄の価格を平均した指標である。本研究では最も広く知られている TOPIX に従った投資手法をベンチマークとして採用する。ある特定の日 (t 日) に TOPIX に従って投資した場合の収益 $\mathsf{topixReturn}(t)$ は、t 日の TOPIX の値 $\mathsf{topix}(t)$ を用いて式 (1) で表される。

¶ 株価と発行株式数の積、企業価値を表す。
‖ 認知度、信頼度共に高い水準の企業が上場する市場を指す。

$$\text{topixReturn}(t) = \frac{\text{topix}(t) - \text{topix}(t-1)}{\text{topix}(t-1)} \tag{1}$$

TOPIX に従い投資を行う事のマクロ的意味は東証市場に上場している全銘柄に平均的な分散投資を行うという事である。従って式 (1) で算出される収益率は、景気の変動と捉えることが出来ることに留意されたい。

2.3 銘柄別投資の収益率と超過収益率 (Abnormal Return)

本研究では株式売買を日々の終了価格においてのみ行うという制約を課しているため、第 t 日の銘柄 x の収益率 $\text{Return}(t)$ は、式 (2) で算出される。$\text{close}(x,t)$ はある銘柄 x の t 日の終値を表している。

$$\text{Return}(x,t) = \frac{\text{close}(x,t) - \text{close}(x,t-1)}{\text{close}(x,t-1)} \tag{2}$$

銘柄別投資の収益率 (式 (2)) と TOPIX 分散投資の収益率 (式 (1)) から、市場の浮き沈みの影響を排除した (TOPIX というベンチマークに対する相対的な) 収益率を超過収益率として定義する。第 t 日の銘柄 x の超過収益率 $\text{aReturn}(t)$ は、式 (3) で算出される。

$$\text{aReturn}(x,t) = \text{Return}(x,t) - \text{topixReturn}(t)$$

$$= \frac{\text{close}(x,t)}{\text{close}(x,t-1)} - \frac{\text{topix}(t)}{\text{topix}(t-1)} \tag{3}$$

以降、収益率という表記が度々出てくるが、特に断らない限り超過収益率の事を意味する。市場に対する相対的な収益率を使って、全ての実験結果を評価する。

2.4 一般的なリスク評価とシャープ比 (Sharpe Ratio)

これまでで、収益率の導出方法についての準備が終わった。ところで、新規の投資手法を評価する際に収益率が重要視されるのと同等に、そのリスク**も重要視される。例えば、同期間に全く同じ収益率を示す投資手法があったとする。この際どちらの投資手法が優秀なのかを決めるならば、収益率の途中経過を参照する事によりリスク評価をする必要がある。安定して収益率が上昇したのか、激しく乱高下を繰り返した後に高い収益率を示したのかでは、全く性質が異なる。前者は優秀な投資手法と言えるが、後者はギャンブルの結果勝利したに過ぎない。本研究ではリスク評価にシャープ比 (Shape Ratio) を採用している。シャープ比とは、ある投資手法における収益率がそのリスクに占める比率を指し、リスクに見合うだけの収益を挙げているか否かの指標である。利回りの同じ投資手法では、シャープ比が高いほど優秀だと言える。ある銘柄に対して、一定期間の投資を仮定すると、シャープ比が理解しやすい。投資期間 k において、銘柄 x に対するシャープ比 $\text{Sharpe}(x,k)$ は、銘柄 x の収益率平均 $\text{ave\_aReturn}(x,k)$ とある銘柄 x の収益率の標準偏差 $\text{usd}(x,k)$ から、式 (4) で導出される。

$$\text{ave\_aReturn}(x,k) = \frac{\sum^{k} \text{aReturn}(x,t)}{k}$$

$$\text{usd}(x,k) = \sqrt{\frac{\sum^{k}(\text{aReturn}(x,t) - \text{ave\_aReturn}(x,k))^2}{k}}$$

$$\text{Sharpe}(x,k) = \frac{\text{ave\_aReturn}(x,k)}{\text{usd}(x,k)} \tag{4}$$

2.5 曜日番号と週番号

本研究では曜日傾向を抽出するにあたり、週単位で統一的に統計処理を行う。これを円滑に進めるため、扱う株価データの日付 (2001 年 1 月〜2015 年 5 月) を週番号 (第 0 週〜第 750 週) と曜日番号 (0〜4 曜日) に変換する。この際、第 0 週は 2001 年 1 月 1 日を含む週としている (表 1)。

週表記	月	火	水	木	金	土	日
曜日番号	0	1	2	3	4	nil	nil

表 1: 任意の 1 週間の曜日番号変換表

例えば、2012 年 11 月 13 日 (火) は第 619 週の 1 曜日として扱う。

2.6 売買パタンの定義

曜日傾向を確認するためには、週単位での売り買いの収益率を網羅的に探索する必要がある。本研究では、第 n 週に対して考慮すべき、銘柄 x の全 20 通りの売買パタンを定義する (表 2)。

ここで強調しておきたいのは、同一曜日に売買を行うパタンは考慮に値しないという事である。例えば、第 n 週の火曜日に購入し第 $n+1$ 週の火曜日に売却をするパタンは火曜日の終値で買って火曜日の終値で売却することになるので、曜日に関係なく常時保有する場合と変わらない。以上のことを踏まえ、網羅的に売買パタンを列挙すると上記の 20 通りとなる。

3. 投資実験

株式データに一連の処理を行い、投資モデルを構築するまでの基本手法について述べる。本研究は曜日に見られる傾向を分析することが目的であるため、週単位の取引に限定した投資手法を網羅的に取り扱っていく。(表 3)
以下 3 つのセクションが、実験の基本構造となる。

** 価格変動が激しいにも関わらず、収益が振るわない場合をリスク高とする。

週番号	n	n	n	n	n	$n+1$	$n+1$	$n+1$	$n+1$	
曜日	月	火	水	木	金	月	火	水	木	収益率
pattern0	+	-								$R_0(x,n)$
pattern1	+	*	-							$R_1(x,n)$
pattern2	+	*	*	-						$R_2(x,n)$
pattern3	+	*	*	*	-					$R_3(x,n)$
pattern4		+	-							$R_4(x,n)$
pattern5		+	*	-						$R_5(x,n)$
pattern6		+	*	*	-					$R_6(x,n)$
pattern7		+	*	*	*	-				$R_7(x,n)$
pattern8			+	-						$R_8(x,n)$
pattern9			+	*	-					$R_9(x,n)$
pattern10			+	*	*	-				$R_{10}(x,n)$
pattern11			+	*	*	*	-			$R_{11}(x,n)$
pattern12				+	-					$R_{12}(x,n)$
pattern13				+	*	-				$R_{13}(x,n)$
pattern14				+	*	*	-			$R_{14}(x,n)$
pattern15				+	*	*	*	-		$R_{15}(x,n)$
pattern16					+	-				$R_{16}(x,n)$
pattern17					+	*	-			$R_{17}(x,n)$
pattern18					+	*	*	-		$R_{18}(x,n)$
pattern19					+	*	*	*	-	$R_{19}(x,n)$

表 2: 第 n 週において考慮される売買パタン
+ : 株式購入日, * : 保持期間, - : 株式売却日

3.1 : 入力データの整形
3.2, 3.3 : 週ごとに動的なポートフォリオの構築
3.4 : テストデータによるシュミレーション

3.1 データ整形

2001 年 1 月〜2015 年 5 月 (0〜751 週) の各週において、上場銘柄 x 毎に 20 通りの売買パタンに付随する収益率 (表 2 $R_0(x,n)$...$R_{19}(x,n)$) を保持する。銘柄数約 4,000 に対し、20 通りの売買パタンと収益率が存在する。従って 1 週間あたりに保持する収益率の総数は、約 80,000 件 ($4,000*20 = 80,000$) となる。この各週について列挙された収益率がモデル構築の際の入力となる。データ処理においては、一般的な PC でも大規模なデータ処理を可能にする"NYSOL"というオープンソフトウェア [7] を使用した。

3.2 ポートフォリオ構築 (モデリング)

本研究において、第 n 週のポートフォリオを構築する事は、買う銘柄群 (ticker コード 1301〜9999 の約 4,000 件) とその売買パタン (表 2. pattern0〜pattern19) のセットを決定することと同意である。この際、過去 k 週分の全銘柄の全売買パタンの収益率を学習に使用する。

学習に使用する週 (k 週分)	ポートフォリオ構築週
$n-k$ ... $n-2$, $n-1$	第 n 週

表 3: 学習の過去参照週数とポートフォリオ構築週との対応表

3.3 リスクを考慮した各銘柄の売買判定

乱高下しているものよりも、一定の推移を保っている銘柄を選択していきたい。リスクに占める収益率の割合を見るためにシャープ比でのフィルタリング判定を行う。そのためには 2.4 節で述べた一般的な指標を拡張する必要がある。各銘柄 $\{\ x \mid 1301 \leqslant x \leqslant 9999\ \}$ $item \simeq 4000$ それぞれに二次元配列データ $A^x_{iu}(\ 0 \leqslant i \leqslant 19,\ 0 \leqslant u \leqslant k-1)$ を定義する。学習に用いる週 (l) を行にとり、それぞれの売買パタン $\{\ i \mid 0 \leqslant i \leqslant 19\ \}$ を列とする。各要素には行と列に対応した収益率 $R_i(x,l)$ を格納する。A^x_{iu} には銘柄 x についての学習に使用する k 週間分の全売買パタンの収益率が格納される。つまり全ての銘柄 x について、この二次元配列データが保持されて初めて第 n 週のポートフォリオを構築する準備が整ったことになる。

$$A^x_{iu} = \begin{pmatrix} R_0(x,n-k) & R_0(x,n-k+1) & \dots & R_0(x,n-1) \\ R_1(x,n-k) & R_1(x,n-k+1) & \dots & R_1(x,n-1) \\ R_2(x,n-k) & R_2(x,n-k+1) & \dots & R_2(x,n-1) \\ R_3(x,n-k) & R_3(x,n-k+1) & \dots & R_3(x,n-1) \\ R_4(x,n-k) & R_4(x,n-k+1) & \dots & R_4(x,n-1) \\ R_5(x,n-k) & R_5(x,n-k+1) & \dots & R_5(x,n-1) \\ R_6(x,n-k) & R_6(x,n-k+1) & \dots & R_6(x,n-1) \\ R_7(x,n-k) & R_7(x,n-k+1) & \dots & R_7(x,n-1) \\ R_8(x,n-k) & R_8(x,n-k+1) & \dots & R_8(x,n-1) \\ R_9(x,n-k) & R_9(x,n-k+1) & \dots & R_9(x,n-1) \\ R_{10}(x,n-k) & R_{10}(x,n-k+1) & \dots & R_{10}(x,n-1) \\ R_{11}(x,n-k) & R_{11}(x,n-k+1) & \dots & R_{11}(x,n-1) \\ R_{12}(x,n-k) & R_{12}(x,n-k+1) & \dots & R_{12}(x,n-1) \\ R_{13}(x,n-k) & R_{13}(x,n-k+1) & \dots & R_{13}(x,n-1) \\ R_{14}(x,n-k) & R_{14}(x,n-k+1) & \dots & R_{14}(x,n-1) \\ R_{15}(x,n-k) & R_{15}(x,n-k+1) & \dots & R_{15}(x,n-1) \\ R_{16}(x,n-k) & R_{16}(x,n-k+1) & \dots & R_{16}(x,n-1) \\ R_{17}(x,n-k) & R_{17}(x,n-k+1) & \dots & R_{17}(x,n-1) \\ R_{18}(x,n-k) & R_{18}(x,n-k+1) & \dots & R_{18}(x,n-1) \\ R_{19}(x,n-k) & R_{19}(x,n-k+1) & \dots & R_{19}(x,n-1) \end{pmatrix}$$

例) A^{1301}_{iu} は以下のように表される。
$A^{1301}_{10} = R_1(1301, n-k)$ は第 $n-k$ 週に銘柄 1301 を pattern1 で取引した場合の収益率である。

$$A^{1301}_{iu} = \begin{pmatrix} R_0(1301,n-k) & R_0(1301,n-k+1) & \dots & R_0(1301,n-1) \\ R_1(1301,n-k) & R_1(1301,n-k+1) & \dots & R_1(1301,n-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_{19}(1301,n-k) & R_{19}(1301,n-k+1) & \dots & R_{19}(1301,n-1) \end{pmatrix}$$

A^x_{iu} の各行 i について、すなはち売買パタン i について、$R_i(x,n-k), R_i(x,n-k+1) \dots R_i(x,n-1)$ の平均を $mR_i(x)$ とすると、$mR_i(x)$ は式 (5) で表される。

$$mR_i(x) = \frac{\sum_{u=n-k}^{n-1} R_i(x,u)}{k} = \frac{\sum_{u=n-k}^{n-1} A^x_{iu}}{k} \quad (5)$$

A^x_{iu} の各行 i について、すなはち売買パタン i について、$R_i(x,n-k), R_i(x,n-k+1) \dots R_i(x,n-1)$ のシャープ比を $sR_i(x)$ とすると、$sR_i(x)$ は式 (6) で表される。

$$sR_i(x) = \frac{mR_i(x)}{\sqrt{\frac{\sum_{u=n-k}^{n-1}(R_i(x,u) - mR_i(x))^2}{k}}} \quad (6)$$

上記の処理を全銘柄について繰り返し適応していく。
ここでフィルタリング用の二次元配列データ F^x_{id} を以下のよ

うに定義する。

$$
F^x_{id} = \begin{pmatrix}
mR_0(x) & sR_0(x) \\
mR_1(x) & sR_1(x) \\
mR_2(x) & \mathbf{sR_2(x)} \\
mR_3(x) & sR_3(x) \\
mR_4(x) & \mathbf{sR_4(x)} \\
mR_5(x) & sR_5(x) \\
mR_6(x) & \mathbf{sR_6(x)} \\
mR_7(x) & \mathbf{sR_7(x)} \\
mR_8(x) & sR_8(x) \\
mR_9(x) & sR_9(x) \\
mR_{10}(x) & \mathbf{sR_{10}(x)} \\
mR_{11}(x) & sR_{11}(x) \\
mR_{12}(x) & sR_{12}(x) \\
mR_{13}(x) & sR_{13}(x) \\
mR_{14}(x) & sR_{14}(x) \\
mR_{15}(x) & \mathbf{sR_{15}(x)} \\
mR_{16}(x) & sR_{16}(x) \\
mR_{17}(x) & \mathbf{sR_{17}(x)} \\
mR_{18}(x) & sR_{18}(x) \\
mR_{19}(x) & sR_{19}(x)
\end{pmatrix}
$$

本実験においては、F^x_{id}から一定のシャープ比 $sR_i(x)$ を下回る行 i を除外することをリスクフィルタリングと定義している。説明する際の便宜上、$sR_i(x)$ の要素の太字部分のシャープ比の閾値未満であったとして、議論を進める。これらの要素を除外した二次元配列データ G^x_{id}は以下のようになる。

$$
G^x_{id} = \begin{pmatrix}
mR_0(x) & sR_0(x) \\
mR_1(x) & sR_1(x) \\
mR_3(x) & sR_3(x) \\
mR_5(x) & sR_5(x) \\
mR_8(x) & sR_8(x) \\
mR_9(x) & sR_9(x) \\
mR_{11}(x) & sR_{11}(x) \\
mR_{12}(x) & sR_{12}(x) \\
mR_{13}(x) & sR_{13}(x) \\
mR_{14}(x) & sR_{14}(x) \\
mR_{16}(x) & sR_{16}(x) \\
mR_{18}(x) & sR_{18}(x) \\
mR_{19}(x) & sR_{19}(x)
\end{pmatrix}
$$

次に、G^x_{id} の中から、最も値の大きい $mR_i(x)$ のインデックス i を $pat(n, x)$ として求める。

$$
pat(n, x) = \arg\max_i \ \{\ mR_i(x) \mid mR_i(x) \in G^x_{id}\ \} \quad (7)
$$

$pat(n, x)$ は第 n 週の銘柄 x の売買パタンということになる。
第 n 週に売買を行う銘柄 x について A^x_{iu} をから $pat(x)$ を求めるまでの一連の処理関数を $pat(n, x) = model(n, A^x_{iu})$‖と定義する。よって、第 n 週のポートフォリオ $Portfolio(n)$ は以下のアルゴリズムに従って構築される (Algorithm 1)。

‖ $pat(n, x) = nil$ の場合, すなはちシャープレシオの閾値を満たす売買パタンが過去 k 週間において存在しない場合は、第 n 週において銘柄 x は取引されないことになる。

Algorithm 1 Construct $portfolio(n)$

Require: $pat(n, x) \notin portfolio(n)$ if $pat(n, x) = nil$
  while $1301 \leqslant x \leqslant 9999$ **do**
    if x *exists in the stock market* **then**
      $pat(n, x) = model(n, A^x_{iu})$
      $pat(n, x)$ to be added $portfolio(n)$
    end if
  end while
  return $portfolio(n)$

$$
portfolio(n) = \begin{pmatrix}
1301 & pattern0 \\
1331 & pattern5 \\
\vdots & \vdots \\
9999 & pattern19
\end{pmatrix}
$$

シュミレーションではテストデータに基づいて、このポートフォリオに従い、投資実験を行う。以上で、ある週 n におけるポートフォリオ構築までの定式化とする。$portfolio_{i0}(n)$ には第 n 週に取引する銘柄番号が、$portfolio_{i1}(n)$ には過去 k 週で学習し、リスクフィルタリングした結果選択された売買パタンが格納されている。

3.4 シュミレーション

全ての $n(k + 1 \leqslant n \leqslant 750)$ について $portfolio(n)$ を算出する。$portfolio(n)$ 自体は第 n 週における銘柄の売買の方法までしか保持しておらず、その際の収益率を見るにはテストデータとの結合が必要である。ここでテスト用の二次元配列データ T^n を以下の様に定義する。第 n 週のポートフォリオ $portfolio(n)$ における時価総額加重収益率 $pfReturn(n)$ と単純平均収益率 $sfReturn(n)$ 算出を経て、(過去参照週数, シャープ比下限値) の組み合わせに対応する戦略全体の加重平均総収益率 $allReturn(k, su)$ と単純平均総収益率 $sllReturn(k, su)$ の算出までのアルゴリズムを示す (Algorithm 2)。$mv(x, n)$ は第 n 週の銘柄 x の時価総額である。

$$
T^n = \left(\begin{array}{ccc|c}
1301 & mv(1301, n) & pattern0 & R_0(1301, n) \\
1331 & mv(1331, n) & pattern0 & R_0(1331, n) \\
\vdots & \vdots & \vdots & \vdots \\
9999 & mv(9999, n) & pattern0 & R_0(9999, n) \\
1301 & mv(1301, n) & pattern1 & R_1(1301, n) \\
1331 & mv(1331, n) & pattern1 & R_1(1331, n) \\
\vdots & \vdots & \vdots & \vdots \\
9999 & mv(9999, n) & pattern1 & R_1(9999, n) \\
\vdots & \vdots & \vdots & \vdots \\
1301 & mv(1301, n) & pattern19 & R_{19}(1301, n) \\
1331 & mv(1331, n) & pattern19 & R_{19}(1331, n) \\
\vdots & \vdots & \vdots & \vdots \\
9999 & mv(9999, n) & pattern19 & R_{19}(9999, n)
\end{array}\right)
$$

$$
portfolio^*(n) = \left(\begin{array}{cc|cc}
1301 & pattern0 & R_0(1301, n) & mv(1301, n) \\
1331 & pattern5 & R_5(1331, n) & mv(1331, n) \\
\vdots & \vdots & \vdots & \vdots \\
9999 & pattern19 & R_{19}(9999, n) & mv(9999, n)
\end{array}\right)
$$

Algorithm 2 Calculate $\mathsf{pfReturn}(n)$, $\mathsf{sfReturn}(n)$

$$
\begin{array}{l}
\textbf{while } k+1 \leqslant n \leqslant 750 \textbf{ do} \\
\quad \textbf{while } 0 \leqslant i \leqslant portfolio(n).size \textbf{ do} \\
\quad\quad Combi(n,i) = (portfolio_{i0}(n), portfolio_{i2}(n)) \\
\quad\quad \textbf{if } Combi(n,i) \text{ exists in } T^n \textbf{ then} \\
\quad\quad\quad portfolio_{i2}(n) = T^n_{03} \\
\quad\quad\quad portfolio_{i3}(n) = T^n_{01} \\
\quad\quad \textbf{end if} \\
\quad \textbf{end while} \\
\quad portfolio^*(n) = portfolio(n) \\
\quad pfReturn(n) = \dfrac{\sum_{i.size} portfolio^*_{i2}(n) * portfolio^*_{i3}(n)}{\sum_{i.size} portfolio^*_{i3}(n)} \\
\quad sfReturn(n) = \dfrac{\sum_{i.size} portfolio^*_{i2}(n)}{750-k} \\
\quad allReturn(k,su) = allReturn(k,su) + pfReturn(n) \\
\quad sllReturn(k,su) = sllReturn(k,su) + sfReturn(n) \\
\textbf{end while} \\
\textbf{return } \; allReturn(k,su),\ sllReturn(k,su)
\end{array}
$$

例えば、過去参照週数 $k = 50$, シャープ比下限値 $su = 0.1$ の投資戦略における総合収益率は $allReturn(50, 0.1)$ で求められる。Algorithm 2 の中で最も注目すべきは、$pfReturn(n)$ を算出する式である。ポートフォリオ $portfolio^*(n)$ が完成すると、第 n 週に取引する銘柄とその売買パタンと収益率が保持された状態になり、そのポートフォリオ自体の収益率 $pfReturn(n)$ を求める必要がある。この際注意しなくてはならないのは、単純平均と同時に、時価総額による加重平均でも集計する事である。TOPIX 自体は時価総額による加重平均により算出されているため、戦略評価も時価総額による加重平均で行う必要がある。単純平均で収益率を算出すると、時価総額の大きい銘柄 (大型株) と小さい銘柄 (小型株) それぞれからの収益率が同等の重みで算出されてしまう。小型株は値動きが激しいため、高いリスクを伴って見かけの収益が上がる場合が想定される。一般に市場流通量の少ない小型株を多く売買すると、その影響を受けて株価が大きく変動してしまうため、現実的な収益率として評価出来ない。機関投資家などが大型株を中心にポートフォリオを運用することからも分かるように、戦略自体の信頼性を保つため、大型株中心での評価を行う必要がある。ここでは単純平均と加重平均の違いも考察するため、両方で集計をしている。

4. 実験結果の解析とその考察

今までのセクションで一連の投資実験の流れを示した。過去参照週数 k とシャープ比下限値 su の組み合わせを複数設定して、その全てについて同様の実験を適応した。それぞれの集計結果を表 4 に示す。表 4 の数値は TOPIX で平均的に分散投資した収益率に対する分散投資の超過収益率を表している。"単純平均"の表には単純平均にて集計した収益率 $sllReturn(k, su)$ が、"時価総額加重平均"の表には時価総額による加重平均で集計した収益率 $allReturn(k, su)$ が格納されている。前者はパラメータによる偏りが少ない事が見て取れる。後者はシャープ比下限値が高く過去参照週数が大きいほど高い収益率を示しているが、ポートフォリオに選択される銘柄数が極端に少なくなっている事が分かる。単純平均の場合、(k, su) の組み合わせに関わらず、6% から 7% 程度の超過収益率を記録している。しかし時価総額加重平均で集計すると、途端に数値が下がってしまう。ここから、シャープ比下限値が低く過去参照週数の短い戦略では、小型株 (時価総額の低い株) が多くの利益を占めていると考えられる。前節でも断った通り、小型株は流通株式数が少ないため、大量の売買を行うと株価に大きな影響を与えていまい、現実的な投資として機能すると評価は出来ない。一方、時価総額加重平均の $(k = 40, su = 0.0) = 0.191$ などは曜日効果を上手く捉えており、全銘柄の半分以上の銘柄に対する分散投資を約 15 年間継続した結果、超過収益が約 19% を示している。これは年率換算で 1% を超えている事になる。年率換算で見れば $(k = 30, 40, 50, su = 0.9, 1.0)$ の超過収益率は最高で約 150% を示しているが、毎週取引されている平均銘柄数は 10 銘柄未満であるため、これも現実的かつ理想的な戦略という評価は難しい。

4.1 Fama-French three factor model による意味解析

ファーマ-フレンチの 3 ファクターモデル (Fama-French three factor model) とは、株式の期待収益率のクロスセクション構造を記述するモデルである。1993 年にユージン・ファーマとケネス・フレンチらにより発表された [6]。モデルの説明制度に長けている事で広く認知されている。本研究においては、3factors ($Rm - Rf$、SMB、HML) を説明変数、目的変数を $pfReturn(n)$ とし、スリーファクターモデルによる重回帰分析を行なった。

$$
\mathbf{pfReturn(n)} \simeq \alpha + \beta(\mathbf{Rm} - \mathbf{Rf}) + \gamma\mathbf{SMB} + \delta\mathbf{HML} \tag{8}
$$

$Rm - Rf$ は市場全体の加重平均収益率、SMB は上場銘柄を時価総額の大中小で 3 つにグループ分けした際の大小グループの収益率の差分、HML はグロース株とバリュー株の収益率の差分をそれぞれ表す変数である。投資実験の結果に対して、本当に曜日による銘柄選択パタン自体に付加価値があるか否かを根拠づける目的で行うため、これら変数についての詳細は省略する。重要なのは 3factors のどの変数が有意であるのかを評価する事ではなく、回帰式の切片項 α の値と p 値の有意性を評価する事である。p 値の有意性判定については、p 値 < 0.05 を満たした場合に有意と評価している。例えば回帰式の切片項が $\alpha = 0.003\%$ かつ有意である場合、その戦略には曜日による付加価値が働いており、年率換算で $0.03\% \times 52$ 週 $= 1.56\%$ の収益を示したということになる。上記を満たし、ポートフォリオに含まれる銘柄数が十分に確保できている場合に曜日効果と認められる (表 5)。表 5 では、(k, su)=$(40, 0.0), (40, 0.1), (50, 0.0), (60, 0.0) (60, 0.1)$ の 5 つの戦略で曜日効果が示された。過去参照週数が一定数以上あり、シャープ比下限値はかなり弱く設定しなければ、曜日効果は有意に働かないと言える。

	k	su = 10	su = 20	su = 30	su = 40	su = 50	su = 60
単純平均	0.0	**0.712**	**0.755**	**0.683**	**0.678**	**0.686**	**0.711**
	0.1	**0.712**	**0.761**	**0.693**	**0.697**	**0.699**	**0.731**
	0.2	**0.716**	**0.766**	**0.707**	**0.695**	**0.677**	**0.741**
	0.3	**0.717**	**0.787**	**0.711**	**0.661**	**0.681**	**0.783**
	0.4	**0.685**	**0.754**	**0.715**	**0.700**	**0.561**	**0.584**
	0.5	**0.679**	**0.775**	**0.743**	**0.629**	**0.717**	**0.700**
	0.6	**0.657**	**0.882**	**0.768**	**0.893**	**1.371**	**1.080**
	0.7	**0.645**	**0.749**	0.256	**0.769**	**0.573**	**0.579**
	0.8	**0.625**	**0.721**	0.284	**0.643**	**0.576**	0.220
	0.9	**0.663**	**0.704**	0.427	**0.631**	**0.525**	0.241
	1.0	**0.641**	**1.145**	**0.957**	**0.892**	**0.787**	**0.590**
時価総額加重平鈞	0.0	0.062	0.093	0.036	0.191	0.155	0.180
	0.1	0.055	0.100	0.034	0.187	0.131	0.161
	0.2	0.046	0.115	0.037	0.096	0.029	0.074
	0.3	0.048	0.146	-0.002	0.036	-0.006	0.320
	0.4	0.027	0.147	-0.010	-0.022	-0.108	-0.121
	0.5	-0.034	0.143	0.258	0.301	0.030	**0.786**
	0.6	-0.067	0.151	0.113	0.494	**1.101**	**0.652**
	0.7	-0.138	0.143	-0.708	**0.836**	**0.544**	**0.559**
	0.8	-0.160	0.287	0.264	**1.280**	**1.101**	**0.910**
	0.9	-0.103	**0.543**	0.428	**1.144**	**1.512**	**1.280**
	1.0	-0.069	**0.564**	**1.430**	**1.284**	**1.445**	**1.259**
平均銘柄数	0.0	3228	3232	3235	**3235**	**3229**	**3189**
	0.1	3130	3063	2999	**2931**	2858	**2761**
	0.2	2937	2648	2359	2080	1826	1587
	0.3	2624	1973	1430	1013	713	503
	0.4	2205	1242	661	344	183	104
	0.5	1741	674	247	96	44	27
	0.6	1304	327	84	30	18	15
	0.7	939	148	31	14	12	11
	0.8	659	67	14	10	8	9
	0.9	456	32	9	8	8	8
	1.0	315	17	7	7	7	7

表 4: 収益率と平均取引数による戦略比較

	k	su = 10	su = 20	su = 30	su = 40	su = 50	su = 60
切片項 α	0 .0	0.00014	0.00018	0.00013	**0.00033**	**0.00027**	**0.00026**
	0.1	0.00013	0.00019	0.00011	**0.00030**	0.00025	**0.00026**
	0.2	0.00010	0.00022	0.00013	0.00024	0.00011	0.00017
	0.3	0.00012	0.00024	0.00011	0.00021	0.00008	0.00056
	0.4	0.00011	0.00029	0.00014	0.00017	0.00003	0.00013
	0.5	0.00001	0.00035	0.00053	0.00067	0.00012	0.00141
	0.6	-0.00006	0.00029	0.00028	0.00087	0.00183	0.00108
	0.7	-0.00013	0.00024	-0.00070	0.00167	0.00092	0.00086
	0.8	-0.00017	0.00028	0.00063	0.00211	0.00166	0.00141
	0.9	-0.00007	0.00116	0.00052	0.00186	0.00239	0.00212
	1.0	0.00002	0.00141	0.00249	0.00226	0.00253	0.00234
p 値	0.0	0.28245	0.14548	0.28704	**0.00903**	**0.02745**	**0.03325**
	0.1	0.33360	0.13672	0.41540	**0.02663**	0.06463	**0.04925**
	0.2	0.47278	0.12378	0.44168	0.16298	0.53693	0.38362
	0.3	0.43520	0.16799	0.60297	0.41670	0.75214	0.06746
	0.4	0.54156	0.16844	0.63154	0.62173	0.94284	0.79322
	0.5	0.97131	0.18870	0.13991	0.20624	0.86458	0.12214
	0.6	0.79375	0.34565	0.58372	0.27600	0.14122	0.46493
	0.7	0.59344	0.55039	0.40069	0.24344	0.60379	0.63956
	0.8	0.53411	0.58200	0.64400	0.25237	0.39988	0.47636
	0.9	0.82338	0.14135	0.76993	0.34353	0.24587	0.30545
	1.0	0.95700	0.21401	0.22586	0.28152	0.23840	0.28077
平均銘柄数	0.0	3228	3232	3235	**3235**	**3229**	**3189**
	0.1	3130	3063	2999	**2931**	2858	**2761**
	0.2	2937	2648	2359	2080	1826	1587
	0.3	2624	1973	1430	1013	713	503
	0.4	2205	1242	661	344	183	104
	0.5	1741	674	247	96	44	27
	0.6	1304	327	84	30	18	15
	0.7	939	148	31	14	12	11
	0.8	659	67	14	10	8	9
	0.9	456	32	9	8	8	8
	1.0	315	17	7	7	7	7

表 5: 曜日効果判定

4.2 売買パタンの選択率

前セクションでは、曜日効果を示す戦略の存在とその収益についての評価を行なった。ここでは各戦略において、今回扱った 20 通りの売買パタンの選択率から、どの曜日に色濃く曜日効果が見られるかを表 6 で確認する。表 6 の各売買パタンの選択率は、5% で選択されるはず (全 20 通りより) なので、5% 以上の部分が太字となっている。火月$_2$ パタン (火曜のみ保持しない) が約 8% と最も高く、次に 月金$_2$ パタン (週末と月曜は保持しない) が約 6.5% と高い。週をまたぐ売買パタンのうち、月曜に売却するパタンは全て 6% 前後であり、総じて 5% より高い値を示していることが分かる。曜日効果を示さなかった戦略、例えばシャープ比下限値 su の大きい戦略においては、週末のみ保持するパタンが 10% を超える値を示す。従って、曜日効果により安定した収益を出す銘柄がごく少数存在し、市場の始まる月曜に値上がりする傾向が強いと考えられる。また、週末のみ保持しない売買パタンは、シャープ比の大きい銘柄に絞っても曜日効果を示した戦略の選択率と変化は見られなかった。従って、週末に値崩れを起こす銘柄が一定数存在すると考えられる。

5. まとめと今後の課題

時価総額の重み無しの分散投資を行うと、リスクフィルタリングの強度によらず一定の超過収益率が得られるが、これは小型株による収益の占める割合が多く戦略としては認められない。

時価総額による重み付きの分散投資を行う場合、限られたごく少数の銘柄のみで取引を行うと TOPIX 分散投資に比べて 2 倍以上の超過収益率を示すが曜日効果が確認されなかった。

曜日効果を捉えた戦略の中では、過去 40 週分からシャープ比下限値 0.0 で学習した戦略が最も良い超過収益率を示し、TOPIX よりも年率は約 1% 増加した。曜日効果という付加価値をうまく捉えた戦略と、その時に機能する売買パタンの順位を発見出来た。

今後の課題としては、曜日効果を各銘柄ごとに抽出し、その時系列推移によりクラスタリングを行う事が挙げられる。曜日効果を示した戦略で、傾向別に分類された投資戦略を採用すれば、年率収益の増加に貢献できると期待される。

曜日効果を示した戦略												
		パターン No.	0	1	2	3	4	5	6	7	8	9
		購入曜日	月	月	月	月	火	火	火	火	水	水
k	su	売却曜日	火	水	木	金	水	木	金	月$_2$	木	金
40	0.0		0.04498	0.04654	0.03962	**0.06155**	0.03864	0.02972	0.03996	**0.08320**	0.03310	0.03767
40	0.1		0.04182	0.04579	0.03989	**0.06441**	0.03466	0.02859	0.04094	**0.08650**	0.02878	0.03700
50	0.0		0.04265	0.04599	0.03902	**0.06208**	0.03774	0.02869	0.03983	**0.08264**	0.03207	0.03681
60	0.0		0.04103	0.04546	0.03836	0.06289	0.03717	0.02781	0.03995	0.08252	0.03146	0.03593
60	0.1		0.03818	0.04373	0.03818	0.06607	0.03250	0.02642	0.04091	0.08536	0.02731	0.03516
		パターン No.	10	11	12	13	14	15	16	17	18	19
		購入曜日	水	水	木	木	木	木	金	金	金	金
k	su	売却曜日	月$_2$	火$_2$	金$_2$	月$_2$	火$_2$	水$_2$	月$_2$	火$_2$	水$_2$	木$_2$
40	0.0		0.06127	**0.06284**	0.04070	**0.07129**	**0.05058**	**0.05947**	**0.06195**	0.04134	0.04732	0.04825
40	0.1		0.06291	**0.06595**	0.03785	**0.07240**	**0.05223**	**0.06217**	**0.05885**	0.04106	0.04855	0.04964
50	0.0		0.06197	**0.06647**	0.04007	**0.07263**	**0.05150**	**0.06128**	**0.06237**	0.04018	0.04754	0.04850
60	0.0		0.06249	**0.06910**	0.03969	**0.07219**	**0.05201**	**0.06250**	**0.06322**	0.03970	0.04710	0.04943
60	0.1		0.06443	**0.07317**	0.03734	**0.07415**	**0.05383**	**0.06571**	**0.05977**	0.03910	0.04806	0.05063
曜日効果を示さなかった戦略												
		パターン No.	0	1	2	3	4	5	6	7	8	9
		購入曜日	月	月	月	月	火	火	火	火	水	水
k	su	売却曜日	火	水	木	金	水	木	金	月$_2$	木	金
40	1.0		**0.12647**	**0.06743**	**0.05369**	**0.05762**	**0.05797**	0.04138	0.03086	**0.05815**	0.04335	0.03871
50	1.0		**0.13177**	**0.07049**	**0.05677**	**0.06372**	**0.05207**	0.03477	0.03026	**0.05695**	0.04248	0.04154
60	1.0		0.12704	**0.06761**	**0.05125**	**0.06525**	**0.05525**	0.04126	0.03381	**0.06289**	0.04108	0.04198
		パターン No.	10	11	12	13	14	15	16	17	18	19
		購入曜日	水	水	木	木	木	木	金	金	金	金
k	su	売却曜日	月$_2$	火$_2$	金$_2$	月$_2$	火$_2$	水$_2$	月$_2$	火$_2$	水$_2$	木$_2$
40	1.0		0.04977	**0.05012**	0.04620	0.04156	0.02729	0.04335	0.04567	0.03139	0.04103	0.04798
50	1.0		0.04549	0.04944	0.05132	0.04023	0.03045	0.04041	0.04774	0.02989	0.04117	0.04305
60	1.0		0.04744	0.04925	**0.05016**	0.04108	0.02835	0.03999	0.04417	0.02963	0.04035	0.04217

表 6: 戦略における各パタン選択率

謝辞

本研究の一部は JSPS 科研費 15H05711 の助成による。

参考文献

[1] 岡田克彦, 羽室行信, 森田裕之, ”周辺文脈アプローチを利用した新聞記事内容と株価に 関する分析”, 日本オペレーションズ・リサーチ学会秋季研究発表会アブストラクト集 2010, pp.128-129, [2010]

[2] 前川浩基, 中原孝信, 岡田克彦, 羽室行信, ”「言葉」にみる株価収益率の予測可能性について：探索的アプローチによるモデル構築と out of sample の実証テスト”, Business accounting review, 12, pp.117-130, [2013]

[3] 岡田克彦, 羽室行信, Stephane,Cheung, ”銘柄別期間全列挙に基づく季節性アノマリーのマイニング” The 30th Annual Conference of the Japanese Society for Artificial Intelligence, 3L4-OS-16b-3, [2016]

[4] Jaffe, Jeffrey, Randolph, Westerfield, ”The Week End-Effect in Common Stock Returns”, The International Evidence, Journal of Finance, 40, 2, pp.433-454, [1985]

[5] Jilin Zhanga, Yongzeng Laib, Jianghong Linc, ”The day-of-the-Week effects of stock markets in different countries”, Finance Research Letters 20 pp.47-62, [2017]

[6] Fama, Eugene F.; French, Kenneth R. (1993), “Common risk factors in the returns on stocks and bonds”, Journal of Financial Economics 33 (1): 3-56, doi:10.1016/0304-405X(93)90023-5

[7] NYSOL:大規模表形式データ解析のためのオープンソフトウェアツール, http://www.nysol.jp