

# なぜ株価学習モデルは似たような銘柄を選んでしまうのか？

Why do the models learning stock prices choose similar stocks?

水田孝信<sup>1</sup> 小林悟<sup>1</sup> 加藤徳史<sup>1</sup> 下妻友成<sup>1</sup>

Takanobu MIZUTA<sup>1</sup>, Satoru KOBAYASHI<sup>1</sup>, Tokufumi KATO<sup>1</sup>, Tomonari SHIMOTSUMA<sup>1</sup>

<sup>1</sup> スパークス・アセット・マネジメント株式会社

<sup>1</sup> SPARX Asset Management Co., Ltd.

**Abstract:** We examine that the models which are inputted very different factors and are learning same stock returns select similar stocks each other. We make 2 type models, neural network model and linear regression model. In both cases, the predictions of returns are more similar, as learning term is shorter. The reasons may be that the past stock returns which all models learned are same, and there are positive feedback mechanisms in the markets.

## 1 最近の実市場での問題

図1に示すように、2007年8月、定量的なモデルを用いた株式ファンドの騰落率が異常な値を記録した[1]。これは定量モデルの予測結果が似ており保有銘柄が似ていたことと、一部のファンドのパフォーマンス悪化による損切売りが他のファンドのパフォーマンスも毀損し次々と悪化するサイクルがあったことが推測されることを指摘している。つまり、(a)モデルの予測結果が似ていたことと、(b)図2のようなパフォーマンス悪化のサイクルがあったこと、の2つの要因が重なりこのようなことが起きたと推測されている。定量的な株式モデルはバックテストを繰り返し過去の情報において良いパフォーマンスが出るようにチューニングするのが普通である。これは言い換えれば、過去の株価を学習しているモデルであるといえる。

本研究では、株価学習モデルにおいて、(a)モデルの予測結果が似ていたことについて、ニューラルネットワークモデルを用いた比較、重回帰モデルを用いた比較を行った。その結果、入力信号が極めて異なる場合においても、予測結果が似てしまうことが分かった。

なお、(b)パフォーマンス悪化のサイクルについては本論稿の主題ではないが、最近の市場を理解するのに大変重要であると考えられる Positive Feedback

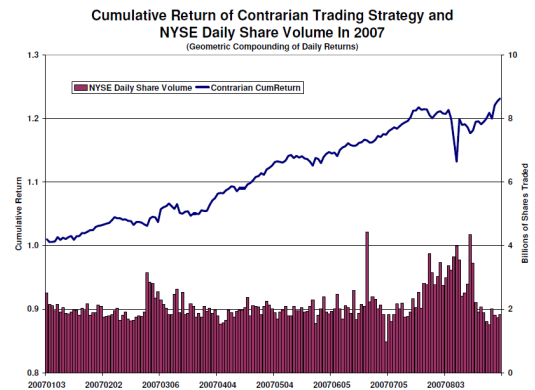
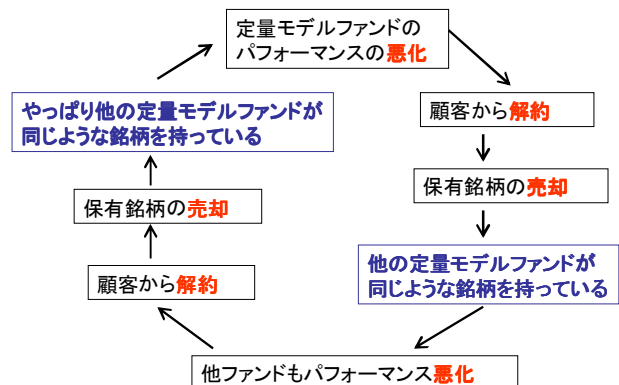


Figure 2: Cumulative return of the contrarian trading strategy from January 3 to August 31, 2007, and the NYSE daily share volume during this same period.

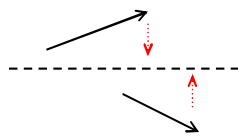
(図1) Khandani and Lo[1]より。2007年8月に定量的モデルによるファンドが異常な騰落を記録した。これは定量モデルの予測結果が似ており保有銘柄が似ていたことと、一部のファンドのパフォーマンス悪化による損切売りが他のファンドのパフォーマンスも毀損し次々と悪化するサイクルがあったことが推測されることを指摘している。



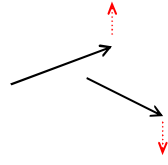
(図2) パフォーマンス悪化のサイクル。いわゆる Positive Feedback を含む系であったと考えられる。

\*連絡先：スパークス・アセット・マネジメント株式会社  
運用調査部  
〒141-0032 東京都品川区大崎 1-11-2  
ゲートシティ大崎イーストタワー16階  
E-mail: [takanobu.mizuta@sparxgroup.com](mailto:takanobu.mizuta@sparxgroup.com)

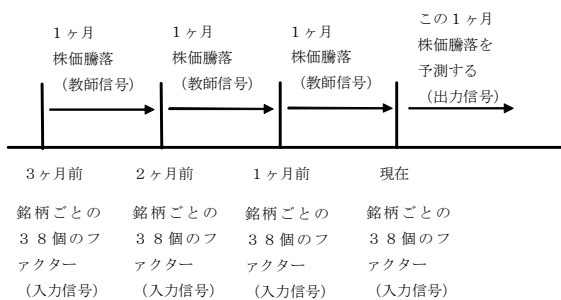
● 古典的なファイナンス・経済理論  
Negative Feedback  
安定的



● 特に最近の市場  
Positive Feedback  
不安定



(図 3) Negative Feedback と Positive Feedback。古典的なファイナンス・経済理論は、実態価格から乖離した価格は Negative Feedback によりすぐに修正され、市場が効率性を保つことを前提にしていることが多い。一方、最近の市場環境は、Positive Feedback を含む系で考えないと説明できない現象が多いと考えられる。



(図 4) 水田ら[5]より。ニューラルネットワークモデルにおいての学習方法。3ヶ月前のファクターを入力信号とし、3ヶ月前から2ヶ月前までのリターンを学習させる。それを2ヶ月前、1ヶ月前についても行う。全ての銘柄に対して学習させたモデルに対し、現在のファクターを入力し、出力されたものを予測の株価騰落率とする。

を含む系である。図3に示すように、古典的なファイナンス・経済理論は、実態価格から乖離した価格は Negative Feedback によりすぐに修正され、市場が効率性を保つことを前提にしていることが多い。一方、最近の市場環境は、Positive Feedback を含む系で考えないと説明できない現象が多いと考えられる。Positive Feedback を含む系の研究は、例えば著名な実務家の分析[2],[3]や多くのシミュレーション研究(例えば[4])がある。しかし、Positive Feedback を含む系の実験やシミュレーションが実際の市場と結びつけられる形で研究されたことはあまり多くなく、今日の市場の理解とあるべき法制度を考える上で、大変重要な研究になると考えられる。

## 2 似た結果を出すモデル

### 2-1 ニューラルネットワークによるモデル

水田ら[5]はニューラルネットワークを用いて予

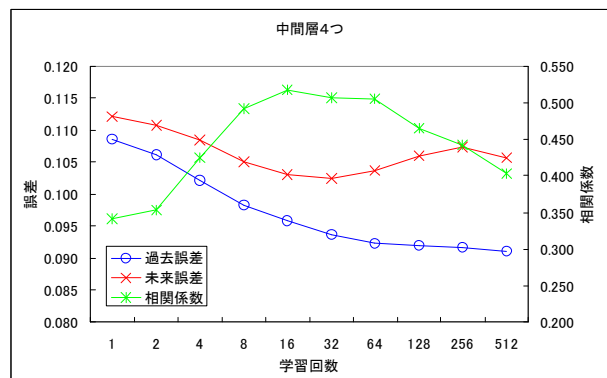
NPMベータ  
時価総額  
BPR  
過去12ヶ月株価騰落率  
過去1ヵ月株価騰落率  
過去3ヵ月株価騰落率  
売買代金/時価総額  
売買代金  
財務レバレッジ  
前期実績ROE  
今期予想ROE  
来期予想ROE  
前期実績EPR  
今期予想EPR  
来期予想EPR  
前期実績配当利回り  
今期予想配当利回り  
来期予想配当利回り

NPMベータ  
時価総額  
BPR  
過去12ヶ月株価騰落率  
前々期～前期 実績売上成長  
前期～今期 予想売上成長  
今期～来期 予想売上成長  
前々期～前期 実績営業利益成長  
前期～今期 予想営業利益成長  
今期～来期 予想営業利益成長  
前々期～前期 実績経常利益成長  
前期～今期 予想経常利益成長  
今期～来期 予想経常利益成長  
前々期～前期 実績純利益成長  
来期予想EPR  
今期～来期 予想純利益成長  
今期売上予想改訂  
来期売上予想改訂  
今期営業利益予想改訂  
来期営業利益予想改訂  
今期経常利益予想改訂  
来期経常利益予想改訂  
今期純利益予想改訂  
来期純利益予想改訂

←  
こちらだけを  
学習させた結果

→  
こちらだけを  
学習させた結果

(図 5)入力信号に使用したファクター。左側と右側のモデルで入力されるファクターの性質は大きく異なる。網掛けされた4ファクターのみは共通。



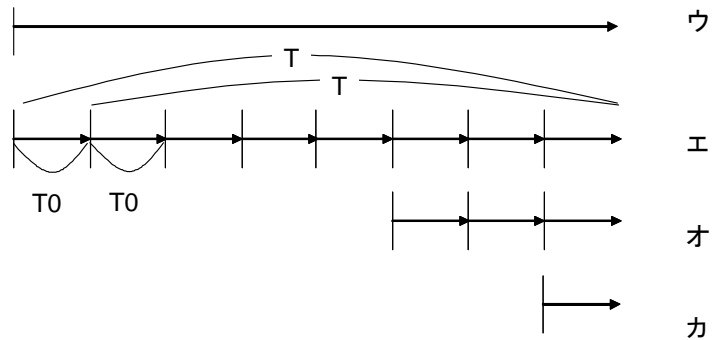
(図 6) ニューラルネットワークモデルの予測結果の類似性。左側のモデルの過去誤差、未来誤差(汎化誤差)と左側と右側の予測結果の相関係数。未来誤差が低いとき、つまり優れたモデルであるときに、最も2つのモデルの予測結果が似てしまう。

測結果の類似性を調べた。図4に学習方法を示した。3ヶ月前のファクターを入力信号とし、3ヶ月前から2ヶ月前までのリターンを学習させる。それを2ヶ月前、1ヶ月前についても行う。全ての銘柄に対して学習させたモデルに対し、現在のファクターを入力し、出力されたものを予測の株価騰落率とする。図5に学習させるファクターを示した。左側と右側のモデルで入力されるファクターの性質は大きく異



Tヶ月前からその後T<sub>0</sub>ヶ月までのリターンを被説明変数とする。

- (ア) 単純平均 (b=1)
- (イ) sgn(b)を係数としT=12, T<sub>0</sub>=12
- (ウ) T=12, T<sub>0</sub>=12
- (エ) T=1~12, T<sub>0</sub>=1
- (オ) T=1~3, T<sub>0</sub>=1
- (カ) T=1, T<sub>0</sub>=1



	相関係数
(ア) 単純平均 (b=1)	-0.09
(イ) sgn(b)を係数としT=12, T <sub>0</sub> =12	0.07
(ウ) T=12, T <sub>0</sub> =12	0.17
(エ) T=1~12, T <sub>0</sub> =1	0.17
(オ) T=1~3, T <sub>0</sub> =1	0.20
(カ) T=1, T <sub>0</sub> =1	0.24

(図 8) 重回帰モデルにおける予測結果の類似性。学習期間を 12ヶ月(エ)、3ヶ月(オ)、1ヶ月(カ)と短くするにつれて相関係数が高まり予測結果がより似てきている。1ヶ月学習の場合、相関係数は0.24とニューラルネットワークモデルほどではないが、入力信号に全く関連性がないこと、重回帰分析という表現能力が低いモデルであることを考えると、かなり高い数字であると考えられる。極めて差別化された作り方をしたつむりの定量モデルでも、予測結果は似がちであることが示されたと考えられる。

な時点  $t-T$  におけるファクター値を説明変数とする。被説明変数  $R_{jt-T, t-T+T_0}$  の数は、 $T$  と  $j$  が可変であるため、 $T=1\sim 12$  と  $j=1\sim 1,000$  から、全部で  $12 \times 1,000=12,000$  個の被説明変数が存在する。過去 12ヶ月の中で1ヶ月ファクターリターンが高いことが多かったファクターのウェイトが高くなる。合成方法(オ)は  $T=1\sim 3$  に縮めたものであり、 $3 \times 1,000=3,000$  個の被説明変数が存在する。合成方法(カ)はさらに  $T=1$  のみに縮めたものであり、1,000個の被説明変数がある。

結果も図8に示してある。合成方法(ア)、(イ)においては、ほとんど相関がない。ところが、過去のファクターリターンを使ってウェイト付けして合成ファクターを作る合成方法(ウ)、(エ)、(オ)、(カ)においては、相関が上昇する。さらに、参照する期間が短いほど、つまり直近のファクターリターンに注目して合成ファクターをウェイト付けすればするほど、相関が上昇している。直近1ヶ月のみ参照にする(カ)が0.24と最も高い相関となっており、左側のファクターと右側のファクターの間で最も相関が高い組み合わせ時の相関である0.200より大きくなっている。ニューラルネットワークモデルほど高くないが、入力信号に全く関連性がないこと、重回帰分

析という表現能力が低いモデルであることを考えると、かなり高い数字であると考えられる。過去のファクターリターンを参照してウェイト付けすることにより合成ファクターが似てしまうことがあることが分かる。

### 3 考察とまとめ

株価学習モデルにおいて、学習信号が著しく異なり、表現空間が狭くない重回帰モデルにおいても予測結果が似てしまうことが分かった。その原因は、学習させる教師信号が株価騰落率という共通のものだからである可能性がある。どんなにモデルに工夫を施しても学習させるものが同じであるため、それにあわせていく過程で似てしまうと考えられる。また、定量ファンドのパフォーマンスが優れていた時期(概ね2001年~2005年)、Positive Feedbackによる”定量モデルファンドのバブル期”においては、モデルが選ぶ銘柄の株価は上昇しがちであったため、株価騰落率自体に定量モデルが選びがちな銘柄という情報が含まれていたと考えられる。つまり、株価騰落率が高い=モデルが選びがち=それを選ぶべき=似せるべきと学習し、似ることが高パフォーマンス

スに直結していたとも考えられる。

いずれにせよ、Positive Feedback を含む系としての市場を実際の市場と対応させて研究することは、最近の市場を理解すること、あるべき規制等を考える上でも大変重要であると考えられ、今後の課題である。

## 参考文献

- [ 1 ] Khandani , A. E., and Lo, A. W.: What happened to the quants in August 2007?, Working paper series, Nov. 4, (2007)
- [ 2 ] Soros, G: The Theory of Reflexivity. New York: Soros Fund Management, (1994)
- [ 3 ] Soros, G: The New Paradigm for Financial Markets: Public Affairs, (2008) ([邦訳] ソロスは警告する: 講談社, 2008)
- [ 4 ] Arthur, W., Holland, J., LeBaron, B., Palmer, R., and Tayler, P.: Asset Pricing Under Endogenous Expectations in an Artificial Stock Market, Technical report, Santa Fe Insitute (1996)
- [ 5 ] 水田孝信, 小林悟, 加藤徳史, 下妻友成 : 精密で複雑なクオンツファンドは優れているか?, 証券アナリストジャーナル, 72-81, 10月号, (2008).