

# 人工知能に関する技術革新と資産運用業の将来： 歴史的概略と経済学的考察

祝 迫 得 夫

## 要 旨

本論文の前半では、機械学習 (machine learning) に代表される近年の人工知能関係のテクノロジーの発展が、資産運用業の最先端の動向に及ぼすについて議論する。株式市場の分析で伝統的に用いられてきたデータセットを用いた取引戦略の構築の過程における機械学習の役割に焦点をあて、経済学・ファイナンスの学術研究における「合理性」の概念と、人工知能による効率的な学習の間にあるギャップの重要性、及びそれがマーケットのあり方に与える潜在的影響について検討する。機械学習の応用が拡大することで、金融取引の自動化・高速化に加え、取引戦略の同質化がこれまで以上に加速することによって株式市場の構造変化がより頻繁になり、マーケットの内在的な不安定性が生み出される可能性を指摘する。論文の後半では、既存のアセット・プライシングの実証研究、特に1990年代初めのFama and Frenchの以降のファクター・モデルの問題点に関して、機械学習のテクニックを応用した最近の研究について議論する。既存のファクター・モデルの研究では、個別銘柄やポートフォリオの収益率の特性を、比較的少数の因子で説明できることを暗黙裡に仮定している。しかし、最近の機械学習を応用した研究では、スパース性と呼ばれるこのような性質が満たされていないことが指摘されている。

キーワード：機械学習、合理性 (rationality)、ルーカス批判、ファクター・モデル、Factor Zoo

## 目 次

1. はじめに
2. AIの技術革新とその資産運用業への応用の歴史
3. 機械学習とファイナンス・経済学
4. 機械学習的アプローチの資産価格モデルの実証分析へのインプリケーション
5. おわりに

## 1. はじめに

「Goldman Sachs はかつて500人いた株式トレーダーを3人に減らした。現在は9,000人のエンジニアを雇用しており、規制を専門とする従業員も増えている。」<sup>1</sup>という、2018年の報道に代表されるように、ここ数年、人工知能(AI)やアルゴリズム・トレーディングが人間のトレーダーを駆逐し、資産運用業の新時代が始まったことを喧伝する記事やシンクタンクのレポートを目にする機会が多い。しかし、新聞・雑誌の報道等で行われているこの種の議論は、事象面においても、アカデミックなテーマ設定という視点からも、複数の本質的に異なるテーマを区別せずにひとまとめにして扱っている。そこでまずは、以下で扱う問題の範囲を明確にすることから始めることにする。

本論文で主なテーマとして扱うのは、人工知能による学習あるいは機械学習(machine learning)と呼ばれる技術の発展が、資産運用業の最先端の動向に及ぼす影響である。これに対し、例えばアルゴリズム取引(algorithm trading)という用語は、機械学習と密接な関係があることには間違いないが、高頻度取引(high-frequency trading)のような、比較的単純なトレーディングをひたすら高速度によって実行することによって利潤を得ようとするような取引戦略も含んでいる<sup>2</sup>。以下では取引の自動化・高速化、取引戦略の同質化といった視点から高頻度取引やファクター投資などにも言及するが、議論の焦点は、取引戦略そのものの構築にあたって機械学習を取り入れている取引戦略／

ビジネス・モデルである。

また、いわゆるビッグ・データ(big data)の利用は、今日のAIを利用した資産運用のビジネス・モデルと切っても切り離せない関係にあり、その範囲はSNSのつぶやきや新聞記事を利用したテキスト・マイニング、小売店などのPOSデータ、記者会見の際の中央銀行総裁の表情の分析、衛星写真による夜間光を利用した景気予測等々、実に多岐にわたる。資産運用における、これらのいわゆるオルタナティブ・データ(alternative data)の利用は近年脚光を浴びているが、その一つ一つについての踏み込んだ議論は技術的なものにならざるを得ず、またどのようなデータを分析対象とするかはあくまで人間の判断となる。以下では主に筆者の専門である株式取引に焦点を絞り、この分野で用いられてきたスタンダードなデータセット、より具体的には様々な頻度の過去の収益率データ、金利変数と、配当・株価比、簿価・株価比、株価・収益比(PER)等の会計関係の変数を利用した、取引戦略の構築・将来予測に分析の焦点を絞る。

本論文の構成は以下のようになっている。第2節では、現在の「第3次AIブーム」以前の人工知能関係の技術のファイナンスの研究と実務への応用について簡単に振り返る。第3節では、経済学・ファイナンスにおける合理性の概念と機械学習における効率的な学習のプロセスを比較しながら、機械学習を金融取引戦略の構築に用いることのインプリケーションについて議論する。第4節では、既存のアセット・プライシングの実証研究、特に近年のファクター・モデルの問題点に関して、最近の機械学習の発

1 Bloomberg (2018)

2 Lewis (2014); 祝迫 (2017)

展から得られた成果を応用した研究について議論する。第5節は、論文全体のまとめである。

## 2. AIの技術革新とその資産運用業への応用の歴史

技術進歩が社会経済に与える影響を考えるにあたって重要なのは、科学技術の発展が直線的なものではないという認識である。松尾豊によれば<sup>3</sup>、これまでのAIに関する技術発展も例外ではなく、重要なブレークスルーと停滞が繰り返されている<sup>4</sup>。そのことを踏まえて松尾は、機械学習、特にディープラーニングと結び付けて議論されることの多い現在の人工知能に関する関心の盛り上がりや、「第3次AIブーム」と呼んでいる。

したがって、マスコミでの報道や社会一般へのインパクトの大きさの違いはあるにせよ、AIに関する技術革新が学問としてのファイナンスや資産運用の実務に重要な影響を与えたのも、今回が初めてではない。例えば、ファイナンスにおける実証研究の発展と実務への影響の拡大において重要な契機となったCampbell達の大学院向けテキストは1997年に出版されているが<sup>5</sup>、その中のNonlinearities in Financial Dataと題した章の一部では、既にニューラル・ネットワークの金融取引戦略の構築への応用が議論されている。また、同章の執筆を担当したAndrew Lo教授は、同時期に統計的学習のテクニックのファイナンスへの応用に関する数多くの研究論文を発表している<sup>6</sup>。

一方、AI技術の資産運用の実務への導入で重要なのは、ジム・シモンズ Sim Simons (元数学者) のメダリオン・ファンド／ルネッサンス・テクノロジーと、デビッド・ショウ David E. Shaw (元コンピューター・サイエンス研究者) のD. E. Shaw & Co.である。これらのファンドについては、かつては具体的な運用手法についてさほど情報が多くなく、またファンドの創業者達もインタビュー等で俗世の脚光を浴びることを避ける行動をとってきた。しかし、少なくともシモンズとメダリオン／ルネッサンスの成功については、ザッカーマン Zuckerman のベストセラーによって、一般にもいろいろなことが知られるようになった<sup>7</sup>。例えば、シモンズのファンドの成功の鍵となっている投資戦略の開発に関わったのが、IBMの音声認識研究グループ出身のエンジニアであること、またシモンズ自身が、自らのファンドで高い収益率を上げているのは、サンプルが多く取れるかなり短期の市場アノマリーを、数量的・統計的に分析した売買に限られるといった認識を述べていることが分かっている。このように、第3次AIブームの実業界への本格的な波及に先駆けて、既に2000年代初頭には機械学習のテクニックがファイナンス／資産運用の実務に取り入れられていた。したがって、次節以降で述べるこの分野での機械学習の限界とみなされている重要な幾つかの問題についても、かなり早い時点で関係者の間では認識されていたのである。

3 松尾 (2015)

4 松尾 (2015)

5 Campbell, Lo, and MacKinlay (1997)

6 Hutchinson, Lo, and Poggio (1994), Lo and MacKinlay (1990), Lo (1994), Lo, Mamaysky, and Wang (2000)などを参照

7 Zuckerman (2019)

### 3. 機械学習とファイナンス・経済学

#### 3.1 機械学習の方法論に関する概観

本節では、主にファイナンス分野で利用する側の立場から、画像認識の例に基づいて機械学習の方法論についてまとめることにする。筆者はAIや機械学習の専門家ではないので、当然のことながらテクノロジーの利用者の立場から見た、かなり雑駁なまとめになる。

多くの動物の画像の中からの「犬」を識別するという作業を考えよう。 $X_i$ をサンプル第*i*番目の画像データ、 $y_i$ をその識別結果だとすると、この作業は以下のような式で表現することができる：

$$y_i \in \{\text{イヌ}, \text{イヌ以外}\} \\ y_i = f(X_i)$$

したがって機械学習の問題は、どのようにして最も効率的な  $f(\cdot)$  を導出するかという問題として記述される。

近年、画像認識のような応用において機械学習の利用が大きく発展した背景には、利用可能なデータ数が大きく増加したこと（いわゆるビッグ・データの利用）や、コンピューターの処理能力の急激な上昇などに加え、深層学習（deep learning）と呼ばれる、より効率的なアルゴリズムの開発・普及があった。実務での応用は多岐に渡り、画像認識やファイナンスの他にも、音声認識、自動翻訳、自動運転、癌の検出による診断支援などの医療分野での応用、タ

クシーの配車予測、ECサイトでの商品のレコメンデーションなど挙げれば切りがない。これらの応用分野に共通するのは、データ数が豊富なことに加え、データの性質がかなり固定であるか、少なくとも大きなデータ構造の変化はごく稀にしか起こらないという点である。

では、機械学習はどのようにして  $f(\cdot)$  を「学習」しているのだろうか？ この問題について議論するために、機械学習の発展のスピードと成果が顕著に現れている例として良く取り上げられる、チェスや囲碁・将棋などの対戦型のゲームについて考えよう。ゲーム理論の言葉で言えば、これらのボードゲームは二人プレイヤーの完全情報のゲーム（two-player games with perfect information）である。このタイプのゲームには、ツェルメロの定理（Zermelo's theorem）と呼ばれる数学的結果が存在することがよく知られている。これは、「もしゲームが引き分けで終わることができないなら、2人のプレイヤーのうちどちらかが必ず勝つ戦略を持つ」という命題である<sup>8</sup>。この結果は、ゲーム理論における「後ろ向き帰納法 backward induction」やナッシュ均衡の考え方の基礎になるものであり、その場合の  $f(\cdot)$  は純粋に論理的な帰結として求められている。

一方、機械学習の方法論は、上記のような純粋に論理的なアプローチとは根本的に異なっている。DeepBlue（チェス）もAlphaGo（囲碁）も基本的には統計的な学習を行っており、膨大な過去の対戦データに基づいて、特定のよく似た局面における最適な指手を統計的な意味で選択している。さらに近年になると、コンピューターの中で仮想的な対戦を繰り返すことで膨大

8 Aumann (1989)

な情報を生産・蓄積し、学習の効率性を飛躍的に高めていると言われる。

これらの議論に関連して、二つの重要な論点を指摘しておきたい。第一に、ミューズ John Muth の有名な論文以降、経済学・ファイナンスで用いられる「合理性」の概念は、経済全体やマーケットの構造をかなり正確に把握している経済主体が、論理的な手続きに基づいて価格や需要・供給に関しての予想を行っていることを意味している<sup>9</sup>。これに対しニューラル・ネットワークや深層学習を利用した投資戦略の構築は、あくまで統計的な意味での最適予測を行っているのであって、Muth 的な強い意味での「合理的」な最適化行動とは本質的に異なっている。問題の設定が非常に明確で利用できるサンプル数が十分に多ければ、ナッシュ均衡のような経済学的な意味で合理的な解と、プレイヤーの統計的な学習の結果として現れる解は、多くの場合一致する。しかし現実の金融市場では、完全に同一なマーケットの状況が繰り返し発生することはないので、経済学的な意味での強い合理性と整合的な価格マーケットで形成されるかどうかは状況次第である。

一般論として、自動化された取引戦略は、緩やかに発生する微小な変化には逐次的な学習・修正で対応できる。しかし、2008年秋のリーマン・ショックとその後の世界金融危機、2020年春からのコロナ感染症の世界的な拡大、2022年に入ってからのロシアのウクライナ侵攻などといった、マーケットに与えるインパクトが非常に大きい、稀にしか起こらないイベントに直面した場合、市場参加者がそれに対応する点に

は時間をかけて学習と修正を行う必要があり、そのことは潜在的に大規模な損失が発生する可能性があることを意味する<sup>10</sup>。したがって、大きなショックに直面した際の新しい投資戦略の見直し・再構築は、AIによる学習を取り入れているか否かに関わらず、経済学の意味での「合理的」な行動というよりは、適応的期待・適応的学習行動 (adaptive expectation/learning) の高度に進化したものとして捉えるべきである。また、現実の金融市場のような断続的に大小のショックが常に発生している環境では、長期に渡って非常に安定的な最適投資戦略は存在しないと考えるべきである。

第二に、機械学習は、利用可能なすべての情報を満遍なく用いている訳ではなく、一部の鍵となる情報を重点的に利用して判断することが可能である点に大きなメリットがある。画像識別の例で言えば、犬の鼻面が猫のそれより長く前に突き出ている傾向にあること、猫の耳は立って尖っているが犬は垂れている傾向にあることなど、犬を特徴づける最も重要な比較的小数の基準をデータから抽出し、それらの基準に基づいて猫を始めとする他の動物との区別をしている。別の言い方をすると、画像ファイルには、犬と他の動物を識別するのに必要な情報以外の余分な情報（ノイズ）が非常に多いので、ノイズを上手く取り除くことができるか否かが非常に重要になってくる。

これは機械学習の分野でスパース性 sparsity と呼ばれる性質であり、高次元で複雑に見えるマーケットのデータも、スパース性を満たしていれば、すなわち共通する特徴を持つ利潤機会

9 Muth (1961)

10 ただし実際に稀な大きなショックが発生したときに、機械学習と人間のどちらが、新しいマーケット環境により迅速に効果的に対応できるかは、また別の問題である。

が繰り返し発生するような環境であれば、機械学習はその効力を存分に発揮できる。この点は、機械学習を取り入れた取引戦略の先駆けである、シモンズのメダリオン・ファンドのようなヘッジファンドが高い収益率を上げているのが、サンプル数が十分に多いイントラデイリーのような取引頻度においてであることと整合的である。逆に、1992年の英ポンド危機の際のジョージ・ソロス George Soros や、サブプライム・バブルの逆売りで莫大な利益をあげたジョン・ポールソン John Paulson のような、非常に儲けは大きい、一回限り、もしくは何十年に一回というような利潤機会に賭けるような投資戦略の構築には、機械学習は馴染まない。また月次・四半期といった経済予測で用いられる頻度のデータが、どの程度スパース性を満たしているかは多分に実証的な問題である<sup>11</sup>。

## 3.2 機械学習と構造変化・ルーカス批判

前節の議論から、機械学習による投資戦略の構築は、マーケットの構造変化がしばしば発生するような状況では、構造変化そのものを統計的にモデル化して予測できない限り、上手く機能しないことが容易に推測できる。

### 3.2.1 マーケットの構造変化としての高頻度取引の登場とフラッシュ・クラッシュ

株式市場におけるそのような構造変化の典型的な例として、2000年代後半の高頻度取引の登場がある。高頻度取引が急速に拡大し、それに特化して利潤を追求する業者が急速に増えたの

は、米国ではリーマン・ショックの前後だったと言われており (Lewis 2014)、日本もそれとあまり間をおかずに、2000年代末までには高頻度取引が大きく拡大している<sup>12</sup>。今日では、注文ベースでは主要な先進国市場のすべての取引の半数以上が、高頻度取引業者によるものとされている<sup>13</sup>。

高頻度取引業者の一部は、他のファンドの注文を察知し、同様の注文をより高速に執行することで先回りして利潤をあげることができていたと言われる。その結果、高頻度取引業者ほど高速ではないが、イントラデイリーでアルゴリズム・トレーディングを行う、既存のクォンツに分類されるタイプの投資戦略の利潤を奪っていたとされる<sup>14</sup>。その意味で、2000年代末以降の高頻度取引の登場は、短期売買の市場に大きな構造変化をもたらし、それ以前に確立されていたイントラデイリーのような比較的高頻度の投資戦略を急速に陳腐化させたと言って良いだろう。

一方で、高頻度取引の普及によって、非常に似通った自動化された取引戦略を採用する取引業者が増え、そのことがマーケットの不安定化をもたらしているという議論が数多くあり<sup>15</sup>、その例として真っ先に挙げられるのが、2010年5月6日のフラッシュ・クラッシュである。ある米運用会社が出したダウ・ジョーンズ平均株価指数先物への大口売り注文をきっかけとした先物価格の急落に、高速・高頻度取引 (HFT) などアルゴリズム取引が追隨して下げ幅を大きくしたとされている。その結果、同一のダウ・

11 この点に関する経済学分野での最近の研究としては、例えば Giannone et al. (2021) を参照。

12 Lewis (2014)、祝迫 (2017)

13 保坂 (2014)、大山他 (2021)

14 Lewis (2014)

15 例えば Stiglitz (2014) を参照

ジョーンズ平均が同日午後の数分間で9% (約1000ドル) 下落し、取引時間中に過去最大の下げ幅を記録した<sup>16</sup>。

一般にフラッシュ・クラッシュは取引の自動化・高速化と、高頻度取引業者というイントラデイリーの取引において大きなシェアを占めるグループの中での取引戦略の同質化が、マーケット全体に潜在的な不安定性をもたらした結果であると解釈されている。ただし、それがマーケット全体のメルトダウンを引き起こすような、致命的な不安定性であるかどうかは定かではない。一方、機械学習はフラッシュ・クラッシュの原因には直接関係していない。しかし、金融取引における機械学習の今後の広がり、取引の自動化・高速化、そして取引戦略の同質化という、フラッシュ・クラッシュの背景にあるマーケットを不安定化させる要因の更なる拡大を引き起こす可能性がある。この点に関しては、3.2.3節で詳しく議論する。

### 3.2.2 取引の自動化・高速化、取引戦略の同質化とクォント・メルトダウン

時期は前後するが、既存のアルゴリズム・トレーディングが大きな損失を被ったとされる他の例として、2007年夏に発生した、クォント・メルトダウン *quant meltdown* と呼ばれている出来事がある。これは2007年8月6日の週に、直前の時期に大きな成功を収めていた、株式ロング・ショート (*long-short equity*) と呼ばれるタイプの投資戦略を採用していたヘッジファンドの多くが、前例のない損失を出した出来事を指す。その背景としては特定のアセット・クラス、具体的には新種の不動産担保証券ないし

はその派生証券のマーケットで発生した混乱があったとされる。これらのアセット・クラスで発生した損失によって、保有する他の資産を急速に流動化させること (*unwinding*) を余儀なくされた金融機関の取引により、ヘッジファンドのリターンを互いに結びつける、事前には観察されていなかった共通の変動要因が出現し、テールリスクを適切に説明・評価できない統計モデルへの過度の依存など他の要因も加わって、ファクター投資モデルにおけるファクター同士の相関、およびファクターと資産価格の相関関係が急激に変化・不安定化した<sup>17</sup>。

つまりクォント・メルトダウンの背後には、具体的なメカニズムは異なるが、取引の自動化・高速化、取引戦略の同質化という、フラッシュ・クラッシュと共通する要因がある。一方で、クォント・メルトダウンによる市場の不安定化は、フラッシュ・クラッシュのように極端な短期間ではないが、それでも1~2週間で終息したとされている。大きな損失が発生した投資戦略は限定的であり、一般的な投資信託などへの影響はほとんどなかった一方で、高頻度取引の登場のケースと同じく、既存のアルゴリズム・トレーディング (の一部) に莫大な損失が発生した。したがってフラッシュ・クラッシュと同じく、クォント・メルトダウンは株式市場に潜在的な不安定性をもたらしたが、それが翌年のリーマン・ショックのような本格的な金融危機に繋がることはなかった。

### 3.2.3 機械学習とルーカス批判

3.2.1節の高頻度取引とフラッシュ・クラッシュ及び3.2.2節でのクォント・メルトダウン

16 Kirilenko *et al.* (2017), 祝迫 (2017)

17 Khandani and Lo (2007)

に関する議論では、金融取引の自動化・高速化、取引戦略の同質化によって、株式市場の構造変化がより頻発するようになり、内在的な不安定性が生み出される可能性について指摘した。これらの問題は、資産運用業における機械学習の利用の拡大によって、より深刻になる可能性があり、そのような議論のロジックは、古くから経済学で議論されてきた、いわゆる「ルーカス批判 Lucas critique」と同じである<sup>18</sup>。

ルーカス批判の最も重要な応用例は、フィリップス曲線として知られるインフレと失業（ないしは産出量ギャップ）の歴史的な負の相関関係と、金融政策の運営に関わるものである。中央銀行が、失業率を恒久的に低下させることを期待してインフレ率を恒久的に上昇させると、やがて民間のインフレ見通しが上昇し、産出量や雇用に関する意思決定が変化してしまう。つまり過去データから推定されたモデルのパラメーターは、政策変更によって人々の行動が変化することで変わってしまう可能性があり、そのようなモデルや分析方法に基づいて政策分析を行うことには根本的な問題があるので避けるべきである、というのがルーカス批判の骨子である。

ルーカス批判のロジックを投資戦略の問題にあてはめると、次のようになる：過去のデータで高い収益をあげている、ある投資戦略があったものとしよう。しかしその戦略が流行して多くの市場参加者が同じような行動を取ると、つまりコピーキャットがマーケットに大量に現れると、それによって分析対象のデータの性質そのものが変わってしまう可能性がある。例え

ば、2000年代までのデータで非常に高い収益性を誇っていた Fama-French の 3 ファクター・モデルのパフォーマンスが、2010年代以降、世界各国で低下し、特に簿価／株価比に基づいたバリューストック投資戦略のパフォーマンスが著しく低下したことは良く知られている<sup>19</sup>。そのような現象の潜在的な説明の一つとして、Fama-French モデルがポピュラーになり過ぎたことで混雑現象が起こり、利潤機会が蕩尽されてしまったという指摘がある。

この種の議論を、より系統的に検証した分析として McLean and Pontiff の研究がある<sup>20</sup>。ある学術論文が、特定の変数による統計的に有意なリターンの予測可能性に関する分析結果を報告しているものとしよう。その予測可能性がミスプライシングを反映しており、学術論文の発表によって洗練された投資家がそのようなミスプライシングについて学習し、それを取り入れた取引戦略を用いるようになると、論文で分析対象となった予測変数と関連する超過リターンは論文発表後に消滅するか、少なくとも大きく減衰すると考えられる。McLean and Pontiff は、クロスセクションの株式リターンのパターンを予測に役立つことが学術論文で報告されている 97 の予測変数について検討を行い、論文の分析が行われた時点以降の out-of-sample のデータ、ならびに論文出版後のデータに基づく予測可能性を検証している。その結果、out-of-sample では 26%、論文公表後では 58% 予測力が低下しており、彼らは後者の結果を、投資家が学術論文からミスプライシング／利潤機会について学習していることの証拠だとしている。

18 Lucas (1976)

19 Fama and French (2020); Cadamuro and Iwaisako (2022)

20 McLean and Pontiff (2016)



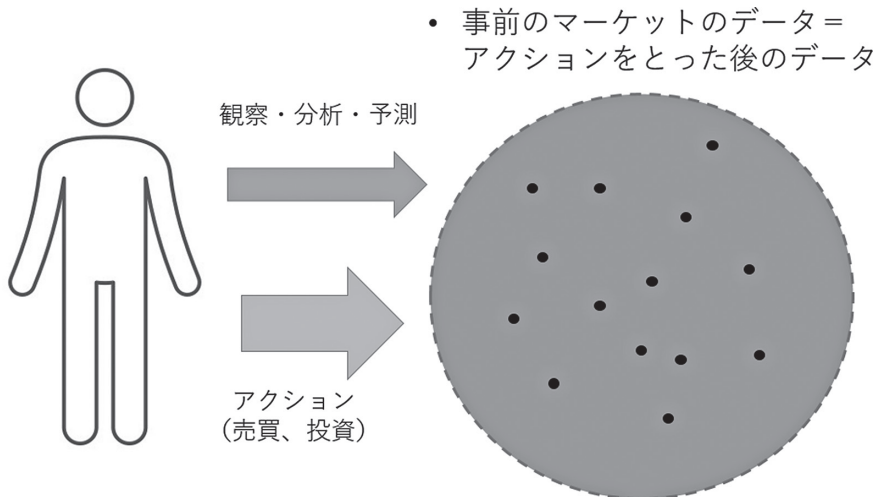
機械学習の典型的な応用例である画像認識と、資産運用業における機械学習の応用を比較した場合、前者では画像認識の結果に基づいて、観察者が何らかのアクションをとったとしても、それによってデータそのものの性質が変化することはない(図表1)。しかし、これがECサイトのプライシングや大規模テーマパークの入場者数予測などの場合、分析対象としている顧客の行動が、観察者のとったアクションによって変化する可能性がある。例えば、大規模テーマパークの入場者数予測がマスコミによって大々的に報道されると、それによって訪れる曜日や時間を変更する人々が多数出てくる可能性がある(図表2)<sup>21</sup>。

資産取引における機械学習の応用が今後さらに進むと、市場参加者の学習はもっと急速なものになっていく。多くの市場参加者が似たような機械学習のテクニックを駆使しており、さら

に類似した予測変数・データセットを使用しているとすれば、あるファンドが過去のデータに基づいて投資戦略を変更する時には、ほぼ同じタイミングで他の不特定多数のファンドが同様な投資戦略の変更を行うと考えられる。その結果、銘柄間の相関関係に大きな構造変化が起こり、過去のデータに基づいて構築された戦略は急速に有効性を失う。本論文で考察しているような、比較的限定されたデータ(過去の収益率データ、金利変数、簿価・株価比等の会計関係の変数)を利用した機械学習のテクニックの利用が今後マーケットで拡大して行くと、分析対象としているデータ、すなわち各変数の平均・分散や共分散/相関などの変数間の関係性の構造が、より頻繁に構造変化を起こすようになる可能性が高い。

さらに極端な状況では、大きなショックに反応した多数のファンドによる一方向の反応が、

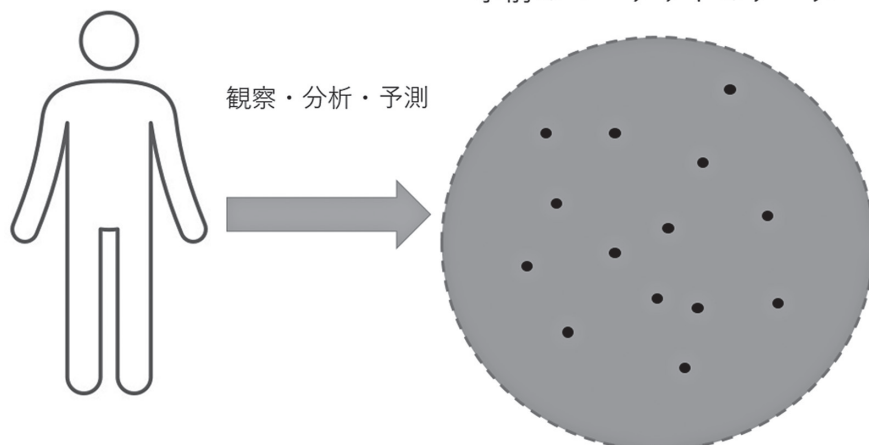
図表1 機械学習が想定する観察者によるデータの分析・予測とアクションの関係



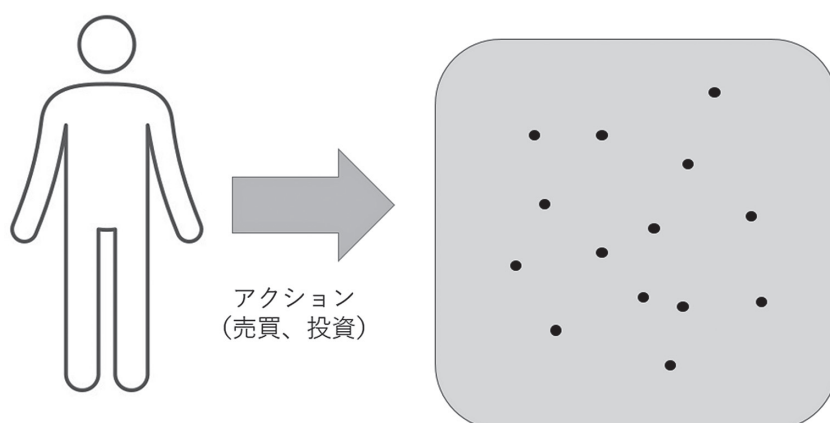
21 このことが明確なのは、電気料金や高速料金におけるピークロード・プライシングの導入である。導入前のデータに基づいて利用時間による利用料金の差別化を行えば、当然利用者の行動パターンは変化する。例えば高速道路料金の深夜割引を導入すれば、トラックなどの大型車の利用は深夜の時間帯へと大きくシフトする。ピークロード・プライシングの場合は、そもそも極端な混雑や遊休状態が発生するのを避け、時間帯による利用状況の平準化を図るために導入するので、そのような構造変化は政策目的になっている。

図表2 ルーカス批判：観察者によるアクションの関係がデータの性質に影響を与える場合

- 事前のマーケットのデータ



- アクションをとった後のデータ



マーケットそのものの不安定化を引き起こす可能性さえある。その典型的な例が、多くの金融機関が同じリスク管理手法を用いていたために、債券価格の下落がマーケット全般での損切りを引き起こし、結果としてさらなる債券価格の大幅な下落を引き起こした2003年のVaRショックである。機械学習が幅広く浸透したマーケットで発生する競争は、定まった対象・データに関して「人間 vs 機械学習」で優劣を競い合うのではなく、「人間 + 機械学習」vs 「人間 + 機械学習」のゼロサム・ゲームにな

る。そのようなマーケット環境下では、過去のデータを前提に導かれた最適な投資戦略は、投資家がそれに基づいて実際のアクションを起こした直後に、陳腐化・非効率化してしまうことになるだろう。

### 3.2.4 その他の問題

投資・資産運用にはその他にも、通常の機械学習の枠組みには必ずしも馴染まない、もしくは想定されていない複雑さがある。第一に、経済学的には資産運用における最適化問題は、将

来の資産から得られる効用の期待値の最大化であり、それは将来の資産水準の期待値と変動の大きさ、両方についての目的関数として表される。しかも分析者は、事後的な資産価格の水準と変動の大きさについてのデータを得るだけであって、直接、投資家の期待値や条件付き分散について観察できるわけではない。したがって、分析対象に関する正しい答えを常に得ることができるわけではないという点において、画像認識のように何が正解であるかについて明確なデータを扱う問題にはない難しさがある。

また現実の投資や資産運用で目標とされるのは、ほとんどの場合、複数の資産からなるポートフォリオのパフォーマンスであって、特定の銘柄／資産のパフォーマンスではない。典型的な例としては、Fama-FrenchモデルのHMLファクターは、簿価・株価比の高い銘柄のポートフォリオと低い銘柄のポートフォリオの収益率の差、つまり前者をロングして後者をショートしたポジションに相当する。したがって投資運用で重要なのは、日経225や円ドルレートの3か月後の値を予想するといった単純な作業ではなく、複数の資産の期待収益率と相関係数の時間を通じた変化について有用な情報を得ることであり、これはより複雑な作業になってくる。

#### 4. 機械学習的アプローチの資産価格モデルの実証分析へのインプリケーション

前節では、機械学習的なアプローチが（主に短期的な）予測に基づく投資戦略の構築において、かなり明確な限界を持っていることを指摘

した。それとは別に、アセット・プライシングにおける既存の学術研究、特に株式収益率のクロスセクションのパターンを、比較的少数の金融・会計関係の変数で説明しようとする、Fama-Frenchの3ファクター・モデルに代表されるファクター投資モデルの枠組みがどこまで有効かという、非常に重要な問題が存在する。主に学術研究の文脈においてではあるが、この問題についてはKozak *et al.* (2020), Nagel (2021)などの最近の幾つかの研究が非常に重要な結果を得ているので、本節ではその問題を取り上げる。

Fama-Frenchの3ファクター・モデルやCarhart (1997)の4ファクター・モデルのような、比較的少数のファクター変数でクロスセクションのパターンを記述できるという前提は、個別銘柄の資産収益率の特性を比較的少数の因子に帰着できるということ、つまりは株式収益率のデータがスパース性を満たすことを暗黙裡に仮定している。しかし、Fama and French以降、非常にパフォーマンスの高いファクター、あるいはファクターの組み合わせに関する論文が多数発表されている。Harvey *et al.* (2016)の整理／メタ分析ではいわゆる共通ファクター (common factors) だけでも113個の、さらに個別銘柄に関する characteristics としては202個の例がとりあげられている。このようなクロスセクションの説明に統計的に有意に役立つ変数の氾濫は、「factor zoo」と呼ばれる奇妙で居心地の悪い状況を生み出している。既に数多くの論文で、このようなあまりにも多数のファクター・モデルの「成功」が孕む問題点について、重要な指摘が行われている<sup>22,23</sup>。

22 Cochrane (2011); Harvey *et al.* (2016); McLean and Pontiff (2016); Feng *et al.* (2020)

23 これは、既に医学分野で“most claimed research findings are false” (Ioannidis 2005) と指摘されていた減少と同じである。

Kozak らによれば、株式市場のデータはス  
パース性を満たしておらず、少数のファクター  
変数を用いたモデルでは、クロスセクションの  
収益率のパターンを適切に要約することは難し  
い。したがって、過去の収益率、会計情報や経  
済情報のデータを利用した分析では人間は AI  
に勝てなくなっていることに加え、Fama-  
French のように、少数のファクターで複雑な  
クロスセクションの期待リターンのパターンを  
説明しようとするパラダイム自体、そもそも無  
理があったという示唆が行われている。

Fama-French 以降のファクター・モデルの  
パラダイムでは、個別銘柄を幾つかの性質／特  
性 (characteristics)、典型的には企業サイズ、  
簿価 / 株価比、事前に推計されたマーケット・  
ベータ等によってソートし、それに基づいてグ  
ループ分けしたポートフォリオのデータを用い  
て資産価格モデルを推計する。実務では、次の  
ステップとして、推計されたパラメーターを用  
いて (しばしばショート・ポジションを含む)  
取引ポートフォリオを構築する訳だが、そのよ  
うにして構築されたポートフォリオのパフォー  
マンスが、完全な等比重ポートフォリオのよう  
な、より単純なポートフォリオと比較して、必  
ずしも高いとは言えないことが良く知られてい  
る<sup>24</sup>。これは古くからある、パラメーターの不  
確実性が存在する下での期待収益率・共分散行  
列の推計と、ポートフォリオ最適化に関する研  
究と深く関連している<sup>25</sup>。つまりパラメーター  
の不確実性が大きな問題になるということは、  
収益率のデータに多くのノイズが存在するとい  
うことであり、したがって、どうやってノイズ  
を削減して、効率的に条件付きの平均と共分散

行列を推計するかという問題になる。このよう  
な視点からは、効率的な学習のあり方に関する  
機械学習の分析フレームワークが、既存のファ  
クター・モデルを含む資産価格モデル一般を、  
包括的に分析・検討するツールを提供してくれ  
ている。

## 5. おわりに

本論文では、現在、様々なビジネスの分野を  
席卷している機械学習の問題を中心に、AI が  
投資・資産運用業で果たす役割について議論を  
行い、まず論文の前半では、主に投資戦略にお  
ける予測の問題に焦点を絞って検討を行った。  
AI が顕著な成功を収めている画像認識等への  
応用と異なり、構造変化の多さや、学習による  
観察者 = 投資家の投資戦略の変更そのものが  
マーケットの構造に影響を与えてしまう可能性  
——いわゆるルーカス批判があてはまる環境——  
などの点において、この分野における AI・機  
械学習の単純な応用には限界があることを示唆  
した。

論文の後半では、1990年代初頭の Fama-French  
モデル以降、30年近くに渡って学術研究・実務  
の両面で大きな影響力を持ってきた、ファク  
ター・モデルの問題点——ないしは「過剰な成  
功」の内包する陥穽——の原因に関する、機械  
学習の視点から見た最近の研究の発展について  
検討を行った。金融市場のデータは元々ノイズ  
が多いことに加え、特に1990年代以降、金融・  
経済関連のデータの利用が飛躍的に容易になっ  
たことによって、潜在的にリターンの予測に役  
立つと考えられる変数があまりにも多くなって

24 Garlappi et al. (2007)

25 Ledoit and Wolf (2004), Brandt et al. (2009), DeMiguel et al. (2009)

しまったことが、話を複雑なものにしている。このような問題の検証においては、機械学習的アプローチによる包括的検証は非常に有効であり、Fama-French以降の資産価格の実証分析において中心的な役割を果たしてきた、比較的少数の金融・経済変数にファクターに基づいたモデルの限界を明らかにし、アカデミックな研究が目指すべき方向を示すことに大きく貢献している。

## 参 考 文 献

- 祝迫得夫 (2017) 「日本における頻度取引 (High Frequency Trading)」の現状について」, 第1期 JSDA キャピタルマーケットフォーラム提出論文 (URL: <https://www.jsda.or.jp/about/kaigi/chousa/JCMF/20160916174151.html>)
- 大山篤之・奥出慎太郎・鈴木賢太・福山義隆 (2021) 「高速取引行為の特性分析」, 金融庁 (URL: [https://www.fsa.go.jp/frtc/report/honbun/2021/20210707\\_SR\\_HFT\\_Article.pdf](https://www.fsa.go.jp/frtc/report/honbun/2021/20210707_SR_HFT_Article.pdf))
- 保坂豪 (2014) 「東京証券取引所における High-Frequency Trading の分析」, 『証券アナリストジャーナル』, 52(6), 73-82.
- 松尾 豊 (2015) 『人工知能は人間を超えるか』, 角川 Epub 選書.
- Aumann, R. J. (1989) *Lectures on Game Theory*, Westview Press. (丸山徹・立石寛訳『ゲーム論の基礎』, 勁草書房)
- Brandt, M.W., Santa-Clara, P., and Valkanov, R. (2009) "Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns," *Review of Financial Studies*, Vol.22, No.9, pp.3411-3447.
- Bloomberg (2018) "A Goldman trading desk that once had 500 people is down to three," May 1st, 2018. (<https://www.bloomberg.com/news/articles/2018-04-30/goldman-trading-desk-that-once-had-500-people-is-down-to-three>)
- Campbell, J.Y., Lo, A.W., and MacKinlay, A.C. (1997) *The Econometrics of Financial Markets*, Princeton University Press. (祝迫得夫・大橋和彦・中村信弘・本多俊毅・和田賢治訳『ファイナンスのための計量分析』, 共立出版)
- Cadamuro L. and Iwaisako, T. (2022) "Value Premium in Japanese Market: Statistical (Re) appraisal," mimeo., Hitotsubashi University.
- Carhart, M.M. (1997), "On Persistence in Mutual Fund Performance", *Journal of Finance* Vol.52, No.1, pp.57-82.
- Cochrane, J.H. (2011) "Presidential Address: Discount Rates," *Journal of Finance*, Vol.66, No.4, pp.1047-1108.
- DeMiguel, V., Garlappi, L., Nogales, F.J., and Uppal, R. (2009) "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms," *Management Science*, Vol.55, No.5, pp.798-812.
- Feng, G., Giglio, S., and Xiu, D. (2020) "Taming the Factor Zoo: A Test of New Factors," *Journal of Finance*, Vol.75, No.3, pp.1327-1370.
- Garlappi, L., Uppal, R., and Wang, T. (2007) "Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach," *Review of Financial Studies*, Vol.20, No.1, pp.41-81.
- Giannone, D., Lenza, M., and Primiceri, G.E. (2021) "Economic predictions with Big Data: the illusion of sparsity," *Econometrica*, Vol.89, No.5, pp.2409-2437.
- Harvey, C. R., Liu, Y., and Zhu, H. (2016) "... and the Cross-section of Expected Returns," *Review of Financial Studies*, Vol.29, No.1, pp.5-68.
- Hutchinson, J.M., Lo, A.W., and Poggio T. (1994) "A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks," *Journal of Finance*, Vol.49 No.3, pp.851-

889.

- Ioannidis, J.P.A. (2005) "Why Most Published Research Findings are False." *PLoS Med* 2(8) : e124. (<https://doi.org/10.1371/journal.pmed.0020124>).
- Khandani, A.E., and Lo, A.W. (2007) "What happened to the quants in August 2007?", *Journal of Investment Management*, Vol.5, No.4, pp.29-78.
- Kirilenko, A., Kyle, A.S., Samadi, M., and Tuzun, T. (2017) "The Flash Crash: High-Frequency Trading in an Electronic Market", *Journal of Finance*, Vol.72, No.3, pp.967-998.
- Kozak, S., Nagel, S., and Santosh, S. (2020) "Shrinking the Cross-section", *Journal of Financial Economics*, Vol.135, No.2, pp.271-292.
- Ledoit, O. and Wolf, M. (2004) "A Well-conditioned Estimator for Large-dimensional Covariance Matrices", *Journal of Multivariate Analysis*, Vol.88, No.2, pp.365-411.
- Lewis, M. (2014) *Flash Boys: A Wall Street Revolt*, W. W. Norton & Company. (渡会圭子・東江一紀訳『フラッシュ・ボーイズ10億分の1秒の男たち』文藝春秋.)
- Lo, A.W., and MacKinlay, A.C. (1990) "Data-snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies* Vol.3, No.3, pp.431-467.
- Lo, A.W. (1994) "Neural Networks and Other Non-parametric Techniques in Economics and Finance," in Fogler H.R. ed, *Blending Quantitative and Traditional Equity Analysis*, pp.25-36, CFA Institute.
- Lo, A.W., Mamaysky, H., and Wang, J. (2000) "Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation," *Journal of Finance*, Vol.55, No.4, pp.1705-1765.
- McLean, R.D. and Pontiff, J. (2016), "Does Academic Research Destroy Stock Return Predictability?" *Journal of Finance*, Vol.71, No.1, pp.5-32.
- Muth, J.F. (1961) "Rational Expectations and the Theory of Price Movements," *Econometrica*, Vol.29, No.3, pp.315-335.
- Nagel, S. (2021) *Machine Learning in Asset Pricing*, Princeton University Press.
- Stiglitz, J. E (2014) "Tapping the Brakes: Are Less Active Markets Safer and Better for the Economy?" paper presented at Federal Reserve Bank of Atlanta, 2014 Financial Markets Conference, Tuning Financial Regulation for Stability and Efficiency. Available from URL: <https://doi.org/10.7916/d8-sg6n-s011>
- Zuckerman, G. (2019) *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution* (水谷淳訳『最も賢い億万長者—数学者シモンズはいかにしてマーケットを解読したか』<上>, <下>ダイヤモンド社)  
(一橋大学経済研究所教授)