

行動ファイナンスと AI による資産運用

関西学院大学 岡田 克彦

伝統的ファイナンス理論では、投資家は合理的であり期待効用最大化させるような投資選択をされると考えられてきました。しかし、ダニエル・カーネマンとエイモス・ツヴェルスキーの提唱したプロスペクト理論の登場後、投資家は必ずしも期待効用を最大化するように判断して取引に従事しているのではないのかと考える研究者が増えてきました。投資家は利得領域においてはリスク回避的に、損失領域においてはリスク愛好的に振る舞う価値関数を持ち、これを最大化しているのではないのかということです。これは株を買ったことがある方であれば誰でも思い当たる節があるのではないのでしょうか。自らが買った株が値上がりした場合、利益をすぐに確定したくなるでしょうし、下落して含み損になった状態では、売却して実現損を出すのではなく、売却を遅らせて株価の回復するのを待ちたくなるでしょう。米国の研究者が、個人投資家がよく使う証券会社の1万口座で調査した結果、果たしてプロスペクト理論と統合的な現象が確認されています。行動ファイナンスでは、わたしたち人間は非合理的な意思決定をするため、これまでの合理性を前提としたモデルでは資産価格を評価できないのではないかと考えるわけです。ところが、伝統的ファイナンスの枠組みにおいては、個人投資家のバイアスはあくまでもノイズに過ぎないと考えます。市場には巨額のお金を動かすプロの投資家が存在し、彼らは合理的にあらゆる情報を鑑みた上で取引価格を決めているはず。したがって、個人投資家がたとえ非合理的に行動しようとも、一時的に誤った価格は「裁定取引」によってすぐに正しい価格に収斂させられると主張します。資金力の大きいプロ投資家が合理的判断をする限り、市場価格はすべての情報を反映した効率的なものになるというわけです。

市場が効率的であるということは、資金調達を行う企業や資産運用をおこなう家計にとってみれば良いことです。市場が効率的であるということは、全ての情報が価格に反映されていることを示します。市場の効率性はファイナンス研究のパラダイムであります。3つの仮説が考えられています。ストロング・フォーム、セミ・ストロング・フォーム、ウィーク・フォームです。最も弱いウィーク・フォームの効率性は、価格には過去に公開情報となった情報が全て反映されていると考えるものです。これは、株価の系列データには何の追加的情報も含まれておらず、過去を研究しても未来はわからないことを示唆しています。もっとも定義の緩いウィーク・フォームにおける効率性は多くの研究者が満たされていると考えていましたが、市場のエビデンスはそれを否定します。その代表的なものが、米国や英国で顕著に観察される「モーメンタム効果」と呼ばれるアノマリーです。これは、過去の一定期間にパフォーマンスの良い銘柄は、近い将来（1ヶ月から6ヶ月）に収益率が相対的に高いという傾向があるというものです。しかも近年だけ観察される現象ではなく、過去1世紀にわたって観察される現象です。情報は株価にしっかり反映されていないのではないのか。行動ファイナンスの立場からの解釈では、投資家は良い情報を受け取ってもすぐには株価に織り込めていない、つまり過小評価していることが原因だと考えます。伝統的ファイナンスでは、リスクが資産価格の収益率を決めるのであり、こうした銘柄選択基準は高いリ

スクを伴うからだと解釈します。リスクか非合理性かの結論は持ち越されたまま、モーメントム効果が報告されて以降は、資産価格評価モデルにもモーメントムファクターが追加されたのです。

現代のファイナンス研究の大きな主流の一つがアセットプライシングモデルです。どのように資産価格が決まっているのかを記述するものですが、業界スタンダードとなっているモデルが 企業規模や成長性を勘案した Fama-French の 3 factor model や先に述べたモーメントム効果を加えた 4 factor model、更に会計情報を活用した Fama-French の 5 factor model などです。こうしたスタンダードなモデルを定規として、研究者たちはモデルでは説明できないアノマリー現象を探索してきました。こうした研究者たちの営みの蓄積は、一流経済雑誌に掲載されたものだけでも現在 250 を超えています。資産価格を評価するのに、250 ものファクターが必要なのでしょうか？250 のアノマリーファクターは同じファクターの別断面を集めたものになっていないのでしょうか？一つ一つのアノマリーファクターはしっかりした検証に基づいていますが、それらを同時に検証することは不可能であり、どの時点で、何が期待リターンを決めているのかはわかりません。これは何とかしなければなりません。2011 年の AFA の会長講演で John Cochrane はこれからのファイナンス研究は、"multi-dimensional challenge"に取り組みなければならないとしています。

私はこの multi-dimensional challenge、つまり次元が多すぎて何をどう整理して良いかわからない現状の、ひとつの解決方法が機械学習を応用することであると考えています。multi-dimensional challenge の一側面は、次元が多すぎて同時検証が出来ない点です。FF3 で扱われている規模と成長性のファクターの 2 つであれば、それぞれの基準で 5 分位に分類したとしても 25 分類ですので、同時検証ができます。しかし、これが 3 つのの基準に増えるだけで 5 分位に分類すれば 125 分類になってしまい同時検証が不可能になります。そこで、ファクターを特徴量ととらえ、それが多いときには階層構造で整理する方が適していると考えます。ファクターという特徴量を使って、買うべき銘柄とそうでない銘柄を決定木で分類するわけです。

機械学習を分類問題に応用する事例は多いです。本講演では、視覚化されていてわかりやすいため、決定木を使った住宅の分類問題を例にとって考えます。今、250 の住宅のデータを見て、それがニューヨークのものか、サンフランシスコのものかを分類させるモデルを考えます。各地域に分類するための特徴量は何が考えられるのでしょうか？例えばサンフランシスコは坂の街ですから、一つは、「海拔」が想起されるでしょう。一方、ニューヨークは物件価格が狭いわりには高いので、もう一つは「平米当たりの単価」も特徴量になります。海拔と平米あたりの単価というデータをつかって、250 の住宅を地域別に分類する学習モデルを作ることができます。もちろん、物件判別にかかわる特徴量が増えれば増えるほど、分類のための情報が増えるわけですので、分類精度が向上します。これを資産運用の文脈で応用することは、住宅を銘柄と置き換え、分類の特徴量をファクターと置き換えて、リターンの高くなりそうな銘柄群とそうでない銘柄群を分類する問題と考えれば良いでしょう。日本の株式で言えば、上場銘柄 3000 社から収益率が高い銘柄群とそうでない銘柄群に分類させるために、特徴量であるファクターを使うということです。AI は決定木をつかったモデルを数多く作りながら、過去をよく説明し、汎化誤差（学習したモデルの out of sample における性能を測る指標）の少ないモデルを選択します。