

アルゴリズム取引の問題点と今後の課題

首都大学東京 大学院経営学研究科 足立高德¹⁾

1 はじめに

コンピュータを使って、つまりプログラムにしたがって自動的に取引を行う手法をアルゴリズム取引といいます。その中でも特に高速に取引を行う手法を高頻度取引あるいは HFT と呼びます。アルゴリズム取引は、完全自動であることが重要です。基本的に人の手を介さず、市場データを観測し、自身で判断し、取引所と通信し、取引を進めていくのです [足立, 2018]。

筆者がアルゴリズム取引を始めたのは、1997年の晩秋でした。戦略記述言語を設計し、その言語で（人間のトレーダーたちが）書いた戦略（プログラム）を解釈して、全自動で取引を行うロボットを開発しました。このとき最初に行った取引は、東証の個別株をロングにし、大証の日経 225 先物をショートにする裁定取引でした。これは HFT に耐えられる十分な流動性を日経 225 先物が持っていたからです。今では大変盛んになっているアルゴリズム取引ですが、当時の東京市場ではまだライバルはいませんでした。このとき、問題だったのは、すべての注文と約定を電子的にやりとりできた東証の個別株と違い、日経 225 先物は、注文こそ電子的に送れましたが、約定がファクスで返ってくることでした。そのため当初はファクスの文字を読んで、端末に打ち込む専門の要員を配置しました。しかしこれでは人の手を介さない全自動とは言えません。そこで当時ようやく使えるようになってきた OCR を使って、ファクスの出力紙を読み取り、自動で電子化するという仕組みを取り入れました。これで一応全自動にはなったわけですが OCR はとても遅く、A4 を 1 枚処理するのに 1 分ほどかかり、人よりも遅いという漫画のような仕掛けでした。それでも、当時としてはこれでも HFT で、十分な利益を生み出していました。その後 1 年ほどして、大証のシステムが変更になり、日経 225 先物の約定も電子的に届くようになり、この不格好な仕掛けを引退させることができました。

本稿では、筆者のこうした経験を踏まえ、先物取引に触れながら、現在と将来のアルゴリズム取引の問題点について考えてみます。

2 取引の高速化に伴う問題点

最初に挙げられる問題点は、微小時間のリスク管理をどうするか、ということでしょう。2010年5月6日にフラッシュ・クラッシュ (flash crash) という事態が発生しました。この日米国東部時間の 14 時 45 分に、S&P500 の先物が、一瞬にして 1000 ポイント (9%) 急落しました。その数分後に暴騰してさらに暴落したのですが、この下手人は HFT ではないかという分析があります [Kirilenko et al., 2017]。一方、直接的な原因は HFT 自体ではなく、投資信託運用会社による超大口の売り注文であり、それをきっかけに普段流動性を供給している HFT が発注を手控えたことが要因ではないか、という見立てもあります [大崎, 2011]。いずれにしても、このようなスパイクが発生するリスクを制御する手法として、戦略の多様化が考えられますが、当時はまだ技術的に現在ほど多種類の戦略を用意するところまでは至っていなかったと思います。

さて、このような問題を抜本的に解決するには、微小時間のリスクを扱う量子リスク理論 (Quantum risk theory) のような道具が必要かもしれません。たとえば時間解像度で言えば、現在は 10^{-3} 秒と 10^{-6} 秒の間

^{*1} Email: tadachi @ tmu.ac.jp Homepage: https://researchmap.jp/taka_adachi

あたりが、HFTの主戦場ですが、これがさらに3桁小さくなり 10^{-6} 秒と 10^{-9} 秒の頻度が当たり前になるのは時間の問題と思われまふ。そこから先はもうSFの世界ですが、もし時間解像度がプランク時間に近づいていったらどうなるでしょう？ そうなるだいで手前で、取引に要するエネルギーが莫大になりすぎて破綻するでしょうが、少なくとも頻度に限界があることはわかります。またそこまで速くなる前に、イベントの前後関係を特定できなくなるでしょう。実際、すでに複数のラインで接続した取引所からの信号は、混雑などの影響により、ラインごとの遅延が異なります。したがって使用する通信路によって観測時刻が異なるのは当たり前になっています。そのため、取引システムのコーディングの際には、観測したイベントの順序が、必ずしも発生順序とは異なる場合も想定して書くことが要求されています。このイベントの観測順序の乱れという問題は、「エネルギー保存則」のような不変性を壊してしまいます。ファイナンス理論で「エネルギー保存則」に相当するのは自己資金充足性 (self-financing property) と呼ばれるもので、時刻 t_0 から t_1 の間に外部との資金の出入りがない、という性質です。ファイナンス理論の様々な定理がこの性質を前提条件として導かれています。高頻度で取引をすると、閉じた系のなかであっても、観測の順序の乱れから、自己資金充足性を仮定した理論が、成立しない瞬間が生じるでしょう。そのような状況から発生するリスクについては、いまだ議論されていないようにみえます。

3 AI取引の問題点

アルゴリズム取引の一形態として、AIや機械学習 (ML) を使った取引があり、すでに先端の技術を使うヘッジファンドや投資銀行は、こうした取引を行っています。本稿の後半では、このAI取引について考察します。まずはAI取引の問題点を、機械学習の問題、確率分布の問題、ブラックボックスの問題の3点についてみていきます。

最初に機械学習の問題です。深層学習に代表される最近の機械学習は、人工ニューラル・ネットワークを多段に組み合わせて実装され、その実行期間は訓練期間とテスト期間に分かれます [巢籠, 2017]。これは、統計学で言う in-sample と out-of-sample に相当します。訓練期間のデータが多くあるほど、経験を積むことができるので、個別株よりも多くのデータ・ポイントを持つ株価指数先物のほうが機械学習に向いていると言えます。

この手法は、前提として、テスト期間に観測されるデータは、訓練期間のそれと同様な統計的性質を持っていることを仮定しています。このため、このような手法を統計的 AI と呼ぶことがあります。こうした統計的 AI の利点はモデル・フリー。つまり特定の確率分布を前提としない点にあります。これは線形回帰法のようにデータにいくつかの性質を要請したりしないので、適用範囲がぐっと広がります。しかしながらテスト期間のレジメが訓練期間のそれとは違ってしまふ、いわゆるレジメ・チェンジが起こると、まったく予想が外れる事態が発生するという欠点があります。さらに、永続的な市場インパクトと同様に、AI の高速な推論はレジメ・チェンジのペースを早めるかもしれません。もしそうなったら、明日に有用な知識は、何も残らないようになってしまふかもしれないのです。このような統計的 AI は、本当に「学習」していると言えるのでしょうか？

つぎに確率分布の問題点をみます。一般に、統計的分布がわかっているれば、検定という手法が使えます。ですが、ニューラル・ネットワークの状態空間はとてつもなく大きく複雑で、たとえばロス関数の適当な分布を見つけることができません。つまりニューラル・ネットワークを使って生成したアルファをもとにしたアルゴリズム取引に使える信頼できるリスク管理手法を発見することは難しいといえます。

さて、金融リスク管理は、歴史的にはつぎの3段階で進んできていると考えられるでしょう。

1. 確率分布がわかったあとでのリスク計測
2. 確率分布が定かでないときのリスク計測 (モデル・リスク)
3. モデルがあまりに複雑で、リスクを計測できない

このうち1番目の手法が古き良き時代、つまり確率分布がわかっていて検定を使えばよい、というもっとも古いリスク管理手法です。それに続いて、リーマン・ショックのときの状況を調べた際には、2番目のリスクが中心課題となりました。いわゆるモデル・リスクです。しかしながら、AIで駆動されたアルゴリズム取引が市場を支配するようになる近未来では、3番目のリスクが最も大きな問題となるように思えます。つまり確率分布がわからない、というリスクです。

そもそも一回しか試行できないマクロな経済現象を確率論で取り扱うことは妥当なのだろうか、という問題が根底にあります。これは金融リスクを確率分布で考えること自体が、ナンセンスなのかもしれないという主張につながるかもしれません。

3点目はブラックボックスの問題です。現在の統計的AIは、特徴量を事前に人間が与えるわけではないので、それが出した判断の背後にある因果関係を探るのが困難なことがしばしばあります。このように因果関係が定かでない世界を金融業界はどこまで許容できるでしょうか？つまり(経済)理論のない、データだけの世界を、誰が信じるのでしょうか。自己勘定で取引をやる分には問題ありませんが、お客様の資金を種銭とするのであれば顧客説明責任の観点から依然障壁としてあります。このようなことを考えると、ともすればAI取引は黒魔術に支配された取引のように捉えられてしまうのではないかと心配になります。

4 アルゴリズム取引の今と今後の課題

アルゴリズム取引が取り込むデータは、時系列方向とクロス・セクショナルな方向の2方向に広がっています。このうち時系列データに基づいて取引を進める手法は、レジメ・チェンジの影響を受けやすく、この手法で代表的なHFTは相対的に取引が減っていくように思われます。一方、クロス・セクショナルなデータは、同時刻に扱える多くのアセットを利用するという意味で、今の、そしてこれからのAI取引の主戦場になるでしょう。その文脈で多様な先物・オプション商品の重要度は今後ますます増していくと思われます。

それでは2節で述べた3つの問題点にはどのようにして対処すればよいのでしょうか？

アルゴリズム取引のシステムを構築していく上で、昔から行われ、現在でも使われ続けている技術に正気テストがあります。これはシステムの鼓動をモニターしてちゃんと動いていることを確認したり、重要なデータを複数のソースから入手し、それらが一致していることをモニターしたりする技術です。この技術を、統計的知見と組み合わせる一般化した技術が異常検知 [井出, 杉山, 2015] で、今後もシステム運用を安定化させる主要な技術として広く使われていくでしょう。

以下では2節の3つの問題点について、ひとつずつその対処法を考えてみます。

レジメ・チェンジによってそれまで学習してきたことが通じなくなるという現象は、統計的AIのもっとも大きな問題でしょう。人間であったならば、レジメ・チェンジが起こっても、それまで経験した現象の背後にあるロジックを理解(あるいは推測)していれば、それを使って新しいレジメでもそれなりの行動を取ることができますが、パターンを統計的に処理しているだけの統計的AIでは、このようなことができません。これを改善するには、人間と同じように論理的な推論機能を組み込む必要があるでしょう。そうした機能を組み込

んだAIのことを推論的AIと呼ぶことがあるようですが、今後研究が進んでいけば、金融にも直ちに応用されるのではないかと思います。

つぎは、確率分布を特定できない、という問題です。現実的には1通りしか生起しない経済事象を、反復試行することによって、大数の法則や中心極限定理を使えるようにするにはどうすればよいかという問題への処方方は、人工的なデータを付け加えてシミュレーションやバックテストを行うことでしょう。このような操作は一般的にデータ拡張 (**data augmentation**) と呼ばれています。ファイナンスにおけるこのようなデータ拡張の方法のひとつとして、人工市場を挙げることができるでしょう [和泉, 2003]。これは計算機上に仮想的な取引所を作り、その上でエージェントと呼ばれる各種役割を持った参加者を模したプログラムに取引をさせる、というものです。これによって、何度でもシミュレーションを繰り返し、人工データを作成し、その統計的性質を調べることができるわけです。こうしたデータ拡張の手法で大切なのは、そうしたデータ生成過程の背後にある理論です。

最後に、ブラックボックスの問題について触れて、この稿を終えることにします。

仕組みのわからないブラックボックスに対処するための伝統的な手法は、リバース・エンジニアリングです。これは聴診器をあてて、システムの中身を推測する手法ですが、AIでも最近これと似たマルチ・タスク学習と呼ばれる手法が使われるようになってきました。これは、統計的エンジンで駆動されている主タスクを、理論をベースにしたいくつかの副タスクでモニターする形の学習システムで、過学習の防止、ひいては汎化性の増大が期待されています。この技術をファイナンスに応用すれば、顧客への説明責任を果たすことなどに利用できるかもしれません [塩野, 2018]。

このようないくつかのサブシステム（あるいは複数のAI）を持つ学習システムを使う場合、サブシステム同士あるいはそれらと人間が会話し、記録するための共通言語をどう設計するか、という問題が将来的には重要になるように思えます。この理論的基礎の候補としては、動的認識論理などがあるでしょう [van Benthem, 2011], [Adachi, 2016]。

参考文献

- [Adachi, 2016] Adachi, T. (2016). [A framework for analyzing stochastic jumps in finance based on belief and knowledge](#). arXiv:1512.00227 [q-fin.MF].
- [Kirilenko et al., 2017] Kirilenko, A., Kyle, A. S., Samadi, M., and Tuzun, T. (2017). [The flash crash: High-frequency trading in an electronic market](#). *Journal of Finance*, 72(3):967–998.
- [van Benthem, 2011] van Benthem, J. (2011). [Logical Dynamics of Information and Interaction](#). Cambridge University Press.
- [足立, 2018] 足立高德 (2018). [アルゴリズム取引](#). 朝倉書店.
- [和泉, 2003] 和泉潔 (2003). [人工市場](#). 森北出版.
- [井出, 杉山, 2015] 井出剛, 杉山将 (2015). [異常検知と変化検知](#). 講談社.
- [大崎, 2011] 大崎貞和 (2011). [フラッシュ・クラッシュから一年](#). *金融ITフォーカス*, 2011(5).
- [塩野, 2018] 塩野剛志 (2018). [アセット・リターン予測 AIとマクロ経済理論の融合](#). 第21回人工知能学会金融情報学研究会.
- [巢籠, 2017] 巢籠悠輔 (2017). [詳解ディーブラーニング](#). マイナビ出版.

本資料に関する著作権は、株式会社大阪取引所にあります。

本資料の一部又は全部を無断で転用、複製することはできません。

本資料の内容は、株式会社大阪取引所の意見・見解を示すものではありません。

本資料は、デリバティブ商品の取引の勧誘を目的としたものではありません。