

令和 6 年 5 月 31 日現在

機関番号：24405

研究種目：挑戦的研究（萌芽）

研究期間：2021～2023

課題番号：21K18439

研究課題名（和文）流動性需給バランス変動のビッグデータ分析に基づく暴落予測法の開発

研究課題名（英文）Crash prediction based on big data analysis of fluctuation in supply and demand balance for liquidity

研究代表者

高田 輝子（Takada, Teruko）

大阪公立大学・大学院経営学研究科・教授

研究者番号：30347504

交付決定額（研究期間全体）：（直接経費） 4,900,000円

研究成果の概要（和文）：暴落の引き金となる売り注文と買い注文のバランスの変動に着目し、ビッグデータの情報量の多さと暴落事象に強い頑健な手法を活用して、株価暴落予測の高精度化を目指した。世界最大のニューヨーク証券取引所の高頻度指値注文統計(NYSE Openbook)の売り手/買い手別情報を用いて流動性需要供給曲線を推計し、その形状の時系列変動や売り圧力・買い圧力の程度に関わる複数のパターンを、まず、明らかにした。次に、データのみに基づいて市場状態を判別して暴落予測・回避を行う学習機械を開発した。これに、市場の「長期トレンド方向」と「市場の一方向化度」の動きを学習させ、暴落の予測と回避の早期化・高精度化を実現した。

研究成果の学術的意義や社会的意義

株価暴落のような異常事象は、深刻な社会的ダメージをもたらすため、その予測の社会的意義は高い。近々の米国株式市場暴落の可能性を憂う声もあり、本研究提案は時宜も得ている。しかし、暴落のような異常事象の統計分析のためには、現代科学が抱える様々な難問の解決を要するため、これまでほとんど行われてこなかった。本研究の行った「データのみからの暴落予測の実現」は、従来法のパラダイムより新しい第4のパラダイムと呼ばれる「データ中心科学」のアプローチにより難題を解決するものであり、学術的意義も高い。

研究成果の概要（英文）：Aiming at achieving higher crash prediction accuracy, this research extracts maximum information about crash from the world biggest limit order book data, NYSE Openbook, by applying robust nonparametric methods advantageous for crash events. First, the liquidity demand-supply curves are nonparametrically estimated based on the statistics by buyer and seller, and several statistical patterns of time series changes and facts related to the degree of buy/sell pressure are found. Next, I developed a machine which learns market conditions to predict future crash. Lastly, the developed machine is extended to learn the newly found facts from the liquidity demand-supply curve analysis; long-term trend direction and the degree of buy/sell pressure, which achieves the earlier crash prediction with higher accuracy.

研究分野：Computational finance

キーワード：流動性 需給バランス 暴落予測 指値 確率密度推定

1. 研究開始当初の背景

株価暴落のような突然の大きな変化は社会・経済に与えるダメージが大きく、そのメカニズム解明・予測は社会的緊急課題である。暴落は多くの因子が絡む複雑なシステムが生み出す事象であるため model misspecification リスクが高く、事前の仮定を置かずにデータのみから事実や変数間の関係を明らかにすることが本来望ましい。しかし暴落のデータ分析は難しい。滅多に起きない暴落はただでさえ情報が少ない上、異常値の多さや確率分布の非正規性(裾厚性・多峰性)等の問題により従来法の適用が難しいからである。ゆえに従来研究はデータを用いぬシミュレーション分析が主流であった。数少ない暴落のデータ分析についても静学的分析か一暴落事例の解析にならざるを得ず、暴落予測の精度も実用レベルには達していない。

株式バブルの発生原因には、過剰期待、過剰流動性、取引制約の変化、などが挙げられるが、株価暴落の直接の引き金となるのは、売り圧力が買い圧力を大幅に上回り、ついには買い手がつかなくなる流動性要因である。

2. 研究の目的

本研究の目的は、これまで困難とされてきた「データのみに基づく株価暴落早期高精度予測法の開発」である。目的達成に向け、以下3つの対応策により困難を打破する。1) 指値注文表全体イメージをノンパラメトリック確率密度推計により可視化し、流動性需給曲線の時系列変動パターンを直接計測し、2) 世界最大株式市場の最長時系列ビッグデータの活用により利用可能情報の少なさの問題に対処し、3) 応募者が開発・選別してきた異常事象に強い手法の適用により、従来法が扱いにくい統計から最大限の情報を効率的に抽出する。

3. 研究の方法

【研究の全体像】

研究は以下の(1)(2)の二方向から行った。

(1) 流動性需要供給曲線の局面別比較分析による暴落前兆の探索

ニューヨーク証券取引所高頻度指値注文統計(NYSE Openbook)の全個別株式指値統計について、買いと売りそれぞれの指値注文の板情報全体イメージのノンパラメトリック確率密度を推計し、更にそれらを市場全体指値価格として、売指値、買指値、それぞれについて、集計する。集計された密度は、当該株式の買い圧力・売り圧力を示すものであり、流動性需要供給曲線を可視化したものと考えられることができる。この流動性需要供給曲線の時系列変動パターンや局面ごとの形状比較により、需給バランスの変化方向や度合の計測や、市場状態の診断や暴落の前兆現象の探索を行った。

(2) 暴落予測法の開発

(1)における発見事実を活用した暴落予測法を開発した。まず、市場状態を学習して暴落リスクを検知し、回避する学習機械を開発した。次に(1)で発見した暴落の前兆現象を、開発した学習機械に学習させ、暴落予測・回避の高性能化を目指した。

【NYSE Openbook : 世界最大最長買手/売手別統計】

指値注文表は投資家がいくらで何株売り(買い)たいかについての注文履歴をまとめたものであり、図1の例のように売り圧力・買い圧力の詳細がわかる。この表の情報は、ビッド・アスク・スプレッド(最高買値と最低売値の差)やマーケット・デプス(最良指値の取引量に基づいて算出)などの流動性の代表値によって捉えられるのが通常であるが、本研究はこの表全体イメージの時系列変動をそのまま分析する。本研究の分析対象のNYSE Openbook は、世界最大規模のニューヨーク証券取引所上場全企業のopen指値注文内容を1/1000秒頻度で記録する利用可能最長時系列の巨大データである。

図1: 指値注文表例

売数量	気配値	買数量
85,000	104	
172,000	103	
101,000	102	
210,000	101	
146,000	100	
	99	47,000
	98	91,000
	97	136,000
	96	88,000
	95	1,200,000

【異常時解析に強いノンパラメトリック確率密度推定法】

株価暴落のような異常時を含む確率密度は、裾厚性や多峰性を持ち、異常値も多いため、従来法の推定精度を大きく損なう。研究代表者はこうした状況における adaptive kernel 確率密度推定法の優位性を比較精度分析により示した(Takada, EJ 2008)。本研究の確率密度推計が関わる箇所は、全て同手法をベースとしている。

4. 研究成果

本研究の各アプローチの成果は、以下のとおりである。

(1) 流動性需要供給曲線の局面別比較分析による暴落前兆の探索

需給バランスに関わる情報を含む事実をいくつか明らかにした。その中から、暴落予測に活用できる統計的パターンとして、以下を選んだ。

- ① 取引価格が長期上昇トレンドのときに売指値のボラティリティーが高まり、長期下落トレンドのときに買指値のボラティリティーが高まる。
- ② 取引価格がピーク圏にさしかかる前に、売指値と買指値の変動の非対称度(市場の一方向化度)が急速に低下し始める。

上記①②を受けて、暴落予測システムの開発においては、長期上昇トレンド局面別に市場状態を捉えた上で、売買指値変動の非対称度を示す指数の情報も入力情報に加えて行うこととした。

(2) 暴落予測法の開発

- ① データのみに基づいて市場局面を判別して暴落予測・回避を行う学習機械の開発：

予測対象の価格時系列のみを入力情報として、長期上昇トレンド局面か長期下落トレンド局面かに注目して市場局面別市場状態を学習し、突然の大きな下落の予測と回避を、従来法より高い性能で行う学習機械を開発した。

図2は、ニューヨーク証券取引所全上場株式の加重平均指数である NYSE 指数について、黒線で示される提案法(LTDC)の下落回避パフォーマンスが、緑線で示されるバイ・アンド・ホールド(対象株式をそのまま保有し続ける)戦略や、他の色で示される主要な従来法を凌駕する様子を示している。縦軸のドロウダウンとは、ピーク時の累積利益からの落ち込み幅であり、投資家など実務家が最も注目するリスク指標である。

図3では、NYSE 指数についての累積リターンの時系列推移が、戦略ごとに比較して示されている。黒線の提案法(LTDC)が最高のパフォーマンスを挙げており、緑線のバイ・アンド・ホールド戦略との差が、価格が大きく下落するときに大きく拡大することが観察される。すなわち、大きな価格下落回避をもたらす影響の大きさが、ここから見て取れる。

この学習機械は、市場状態をモニターして暴落リスクを感知して暴落回避行動をとるものである。ここでは、市場状態の局面を判別する学習において、長期上昇トレンド局面か長期下落トレンド局面かに注目しているが、学習内容を差し替え、もしくは、追加することにより、暴落予測・回避性能を更に改善できる、普遍的なプラットフォームとして利用できるものである。

- ② 市場の一方向化度を用いた暴落予測

上記①の暴落予測システムに、「市場が売りか買いの一方向に傾く程度」を数値指標化したものを学習対象に追加したところ、暴落予測・回避のパフォーマンスを、より早期に、より高性能に改善することを確認できた。

ただ、この結果については、まだ、業績の形で発表できていない。研究期間中の急激な円安を受けた NYSE Openbook 統計購入費の予算大幅超過分を他の RA 雇用費やデータ格納用 HDD 費をはじめとする他の費目で充填した影響で、新規購入データのクリーニング作業やデータ格納環境整備等を独力で行わざるを得なくなり、それに想定以上の時間と労力をかけざるをえなくなってしまう事情が背景にある。そこで、研究期間内には計算結果を出すことを最優先としたため、目標としていた研究結果を得ることはできなかったものの、研究期間内に成果の形で発表することに十分な時間と労力を割くことができなかった。研究期間終了後にはなってしまうが、順次、成果として発表していきたい。

図2: 下落回避性能

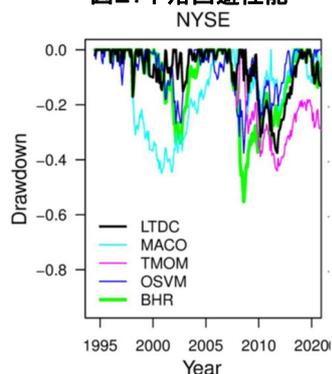
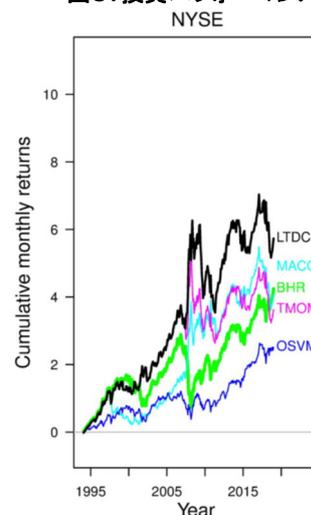


図3: 投資パフォーマンス



<引用文献>

Teruko Takada, "Asymptotic and qualitative performance of nonparametric density estimators: A comparative study," The Econometrics Journal, 11(3), 554-572, 2008.

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 1件/うち国際共著 0件/うちオープンアクセス 1件）

1. 著者名 Takada Teruko, Kitajima Takahiro	4. 巻 17(10)
2. 論文標題 Trend-following with better adaptation to large downside risks	5. 発行年 2022年
3. 雑誌名 PLOS ONE	6. 最初と最後の頁 1~31
掲載論文のDOI（デジタルオブジェクト識別子） 10.1371/journal.pone.0276322	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Teruko Takada and Takahiro Kitajima	4. 巻 7
2. 論文標題 Mitigating the dilemma of trend-following in averaging window selection to better adapt to downside risks	5. 発行年 2021年
3. 雑誌名 OCU Working Paper	6. 最初と最後の頁 1-27
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
---------------------------	-----------------------	----

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------