

機械学習と株価予測

—ニューラル・ネットワークの活用—

金融研究所 三木翔太*、五島圭一

Bank of Japan Review

2020年11月

機械学習を活用した資産価格の予測が盛んに実施されている。なかでも、ニューラル・ネットワークへの期待は大きい。ニューラル・ネットワークは、パターン認識能力に優れた技術で、様々なデータの動きから価格変動の規則性を発見することを可能としている。ただし、予測精度が向上するかどうかは金融市場の情報効率性に依存する。金融市場が効率的であれば、入手可能なあらゆる情報が瞬時に資産価格に反映されるため、予測精度は改善しにくい。実際、近年の日本の株式市場を対象にした簡単な実験によると、標準的なニューラル・ネットワークは、株式収益率の予測精度を向上させないとの結果となる。この結果は、過去の株式収益率のみを用いた予測をベースとしている。予測精度の改善の有無を評価するためには、過去の株価だけでなく幅広い情報を用いるなど一段の検証が必要となる。

はじめに

機械学習の活用が様々な分野で進んでいる。なかでも、ニューラル・ネットワーク (Neural Network: NN) がその中核的な役割を担っている。NN を活用した技術は、既に翻訳や画像認識等の分野で大きな成果をあげており、現在でも、自動運転をはじめ幅広い先端分野で応用されている。

NN の成功の背景には、その優れたパターン認識能力がある。NN は、データから規則や構造を見出すという工程を自動的に行うことを可能としている。例えば、機械翻訳では、NN が原文と訳文の大量の組から尤もらしい翻訳パターンを自動的に見出すことで高い性能を発揮している¹。

資産価格の予測にも NN が盛んに活用されている。従来であれば、人が様々な経済・金融指標を観察し、将来価格を予測するモデルを作成していた。しかし、経済・金融指標と資産価格の間に成り立つ関係には、非線形なものまで含めると膨大な組み合わせが考えられ、人手で設計された予測モデルは有用な関係を見落とす恐れがある。この点、予測誤差を最小にするよう変数や関数形の最適な組み合わせを自動的に選択できる NN には、このような見落としは原理的にはなく²、予測精度の改善への期待が大きい。

しかし、NN も万能ではなく、状況によっては望んでいた性能を発揮しないこともある³。具体的な例として、裏表が同一の確率で出るコイン投げの結果を予測する問題を考えてみよう。コイン投げの過去 10 回の情報を基に、次にコインが表となるか裏となるかを NN で予測しようと試みても、正解率が安定的に 50%を上回ることはないであろう。

将来の資産価格を予測する取り組みも、コイン投げの予測と同様、予測精度の改善に NN が適さない可能性がある。伝統的なファイナンス理論の基礎をなす効率的市場仮説によれば、過去のデータをもとに資産価格を予測することは不可能である⁴。仮に、この仮説が成り立てば、NN のような高精度の予測技術を用いたとしても、予測精度は改善しないと考えられる。

そこで、以下では、NN と効率的市場仮説の関連を探るため、それぞれの概要に触れたあと、標準的な NN を利用した株式収益率の予測に関する簡単な実験を行う。

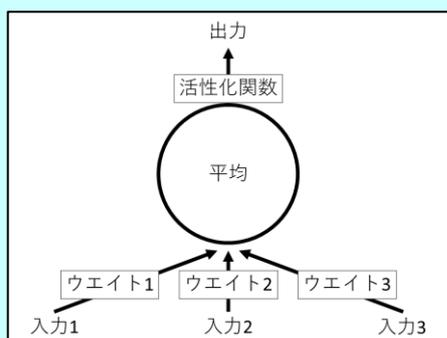
ニューラル・ネットワークの概要

NN は、ヒトの脳の働きの一部を模倣した数理モデルであり、人工知能を実現するための技術とし

て利用されている⁵。脳の中では、ニューロン（神経細胞）が互いに様々な強度で結びついてネットワークを構成している。ニューロンの電気的な興奮が他のニューロンの興奮を惹起し、これが次々に伝播していくことで情報が処理される。ニューロン同士の結合の強さが化学的に更新されることで、情報処理の仕方が徐々に変わり、学習が起こる。

NN を構成するニューロンは⁶、電気的な興奮の代わりに数値（スカラー）をやり取りする。各ニューロンは、他の複数のニューロンから出力された数値を受け取って重み付きの平均値を計算した後、活性化関数と呼ばれる非線形関数を通して出力値を算出する（図表1）。この平均値を計算する際に使われるウエイトがニューロン間の結合強度にあたる。

【図表1】ニューロンの模式図



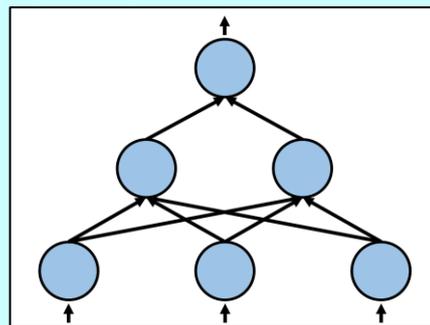
複数のニューロンが層状に積み重なったネットワークを想定することが一般的であり、多くの層を持つNNを用いた学習が深層学習と呼ばれる。一層あたりのニューロン数を増やしたり層を積み重ねたりすることで、あらゆる入出力関係を再現できることが知られている。NNのタイプは、層の重ね方とニューロンの結合の仕方で異なり、それによって捉えやすいデータの構造にも違いがある。以下のとおり、順伝搬型NN、畳み込みNN、再帰型NNの3つのタイプが代表的である⁷。

①順伝播型ニューラル・ネットワーク

順伝播型NN (Feedforward Neural Network: FFNN) は最もシンプルなNNで、各層で受け取った数値情報をニューロンによって処理し、次の層へと順番に伝播していく構造を持つ（図表2）。各ニューロンが、一つ下の層の全てのニューロンからの出

力値を受け取るという特徴を持つことから、全結合型とも呼ばれる。

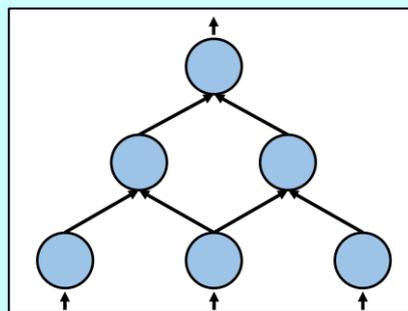
【図表2】順伝播型NNの模式図



②畳み込みニューラル・ネットワーク

畳み込みNN (Convolutional Neural Network: CNN) は、層から層へと順番に情報が伝播していく点は順伝播型NNと同じだが、隣り合うニューロンの情報を順次まとめていくところに特徴がある（図表3）。主に画像処理の分野で多用される⁸。

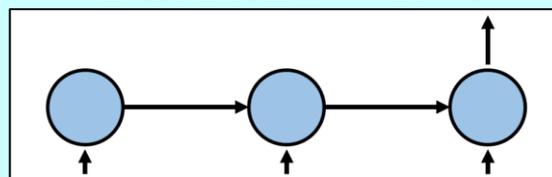
【図表3】畳み込みNNの模式図



③再帰型ニューラル・ネットワーク

再帰型NN (Recurrent Neural Network: RNN) は、入力の順序関係を考慮して処理する構造を持つ（図表4）。時系列データ分析や自然言語処理の分野で多用される。過去の情報をより長い期間記憶できるように改良を施した長短期記憶 (Long-Short Term Memory: LSTM) がよく用いられる。

【図表4】再帰型NNの模式図



一般に、ニューロン間の結合強度を決めるウエイトは膨大な数にのぼり、これらをどのように決

めれば最適な入出力関係を実現できるのかは自明ではない。しかし、ウエイトの値を、データに基づいて自動的に最適化する技術が開発されているため、大量のデータをNNに提示しさえすれば適切な値を求めることができる⁹。

こうした特性は、非線形関数まで含めた柔軟な関係を、人手で明示的に設計することなく、データから自動的に推定することを可能にする。したがって、NNを用いれば、従来の手法では見過ごされていたデータの構造を捉えられる可能性がある。そうであれば、過去の株価の推移を基にNNを用いて明日の株価を予測することによって、これまでにない正確な予測ができ、超過収益を得られる期待が高まる。ただし、この点は、以下に述べる金融市場の情報効率性に依存する。

効率的市場仮説

金融市場は、個別企業の動向、政治的なニュース、マクロ経済指標といった世の中の様々な情報を消化して資産価格に織り込んでおり、新たな情報の内容に応じて資産価格が上下する。例えば、ある企業が買収される際には、その企業の株価は上昇する傾向にある¹⁰。

金融市場による情報の消化が効率的で、利用可能なあらゆる情報が瞬時に資産価格に反映されるとする説が、効率的市場仮説である。言い換えると、あらゆる情報は既に資産価格に織り込まれているため、過去の情報を用いて将来の資産価格を予測し、超過収益をあげることはできないということになる。この仮説は、多くの学者や実務家の関心を引き寄せ、20世紀後半から現在に至るまで盛んに検証が行われている。

効率的市場仮説は、資産価格に織り込まれる情

報集合の大きさに応じて、①ウィーク型、②セミストロング型、③ストロング型の3つのタイプに分類される。ウィーク型は、情報集合として資産価格の現在と過去の値のみを想定している。セミストロング型は、資産価格だけでなく、現在までに全ての市場参加者に知られている統計やニュース等の公的情報を想定している。ストロング型は、これらの情報に加えて、「A社が将来的に買収されるらしい」といった私的情報も含む¹¹。それらの関係を図示すると図表5のようになる。

このうち最も厳しい仮説であるストロング型は、市場参加者の情報収集コスト等を踏まえると成立しないと考えられているが、セミストロング型とウィーク型は十分成り立ち得る。これまでに、これらの仮説を検証するため、時系列の持つ統計的な性質に関する検定や¹²、種々の予測モデルに基づいた売買成績の検証等が多数実施されている¹³。反証はあるものの、ウィーク型およびセミストロング型は概ね成立しているという結果が多い。特に、①金融システムの整備が進んでいる先進国であるほど、②活発に取引される流動性の高い資産であるほど、③短い期間の予測であるほど¹⁴、仮説は当てはまりやすいとされている。さらに、近年の情報処理技術の発展やアルゴリズム取引の増加に伴って、あらゆる情報を瞬時に株価に織り込もうという動きはますます活発化し、市場はより効率的になっている可能性がある。

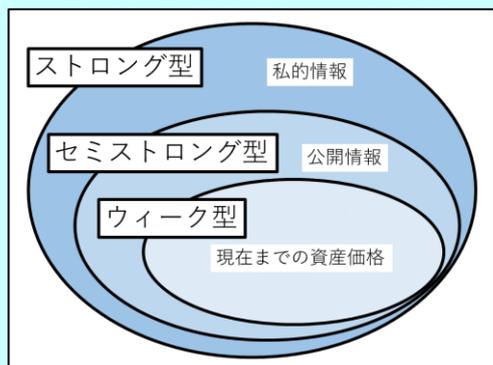
反証としては、標準的な理論では説明できない価格変動の存在（アノマリー）が指摘されているほか¹⁵、国や資産の種類、予測の期間の長さによって結果は変わり得るとの実証結果も報告されている。これらの事実を説明すべく、行動ファイナンス等の新たな学説が生まれるなど、市場の効率性を巡る議論は未だに尽きていない。

次節では、NNと効率的市場仮説の関連を探るため、標準的なNNを用いて、近年の日本の株式インデックスの収益率を予測する簡単な実験を行う。

株式収益率の予測精度比較

効率的市場仮説のもとでは、株式収益率はランダム・ウォークに従い¹⁶、合理的に期待される収益率と事前に予測できない誤差で構成されている。すなわち、株式収益率はその期待値を中心にはら

【図表5】 情報集合の包含関係



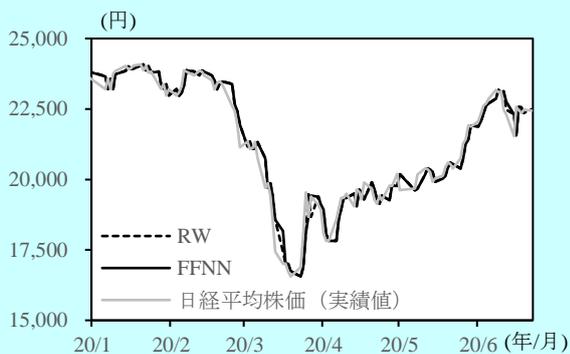
つくであろうことは分かっていても、それ以上の予測を立てようがないということを意味している。

ここでは、①日経平均株価、②日経平均先物、③TOPIX、④TOPIX 先物の4つの指標の日次収益率の予測を試みる。予測は、それぞれの日次収益率の過去の値のみを用いて行う。効率的市場仮説をベースとした予測モデルとして、ランダム・ウォーク・モデル (RW) ¹⁷を用い、同仮説に反するモデルとして、自己回帰移動平均モデル (ARMA) と NN を活用したモデルを使用する。ARMA モデルは、今期の収益率が過去の収益率・誤差と線形の関係にあると仮定したモデルである¹⁸。非線形関係を許容した NN モデルとして、順伝播型 (FFNN)、畳み込み (CNN)、再帰型 (LSTM) の3種類を使用する¹⁹。予測期間は、2017年12月28日から2020年6月22日とした。まず、2017年12月28日の予測値は、2000年2月3日から前日 (2017年12月27日) までのデータを用いてモデルを推定することで算出した。それ以降の予測値は、データの期間を1日ずつ後にずらしながらモデルを推定することで算出した²⁰。

NN や ARMA による収益率の予測精度が仮に RW を上回らないのであれば、金融市場はウィーク型の意味で効率的であると解釈できる。つまり、NN や ARMA の予測を基に超過収益を得る余地がないことを示唆する²¹。

図表6は、日経平均株価の2020年の実績値と予測値 (RW, FFNN) を示している。2つのモデルに基づく予測値は、ほぼ重なって推移しており、モデル間の差はほとんどない。図表7は、モデルごとの予測精度を示している。各数値は予測誤差の大きさを示しており、値が小さいほど予測精度

【図表6】日経平均株価の推移とモデルによる予測値



(注) RW と FFNN は、前日の株価 (実績値) と日次収益率の予測結果から算出。

【図表7】収益率の予測誤差

＜日経平均株価＞

	RW	ARMA	FFNN	CNN	LSTM
MAE	0.009	0.009	0.009	0.009	0.011
RMSE	0.013	0.013	0.013	0.014	0.022

＜日経平均先物＞

	RW	ARMA	FFNN	CNN	LSTM
MAE	0.010	0.010	0.010	0.010	0.011
RMSE	0.014	0.014	0.015	0.016	0.019

＜TOPIX＞

	RW	ARMA	FFNN	CNN	LSTM
MAE	0.008	0.008	0.009	0.008	0.010
RMSE	0.012	0.012	0.016	0.012	0.018

＜TOPIX 先物＞

	RW	ARMA	FFNN	CNN	LSTM
MAE	0.009	0.009	0.010	0.009	0.010
RMSE	0.013	0.013	0.015	0.015	0.017

(注) 2017年12月28日～2020年6月22日における予測誤差の平均値を示している。予測誤差の評価には、株式収益率の予測誤差の文献で一般的に用いられている、平均絶対誤差 (Mean Absolute Error: MAE) および二乗平均平方根誤差 (Root Mean Squared Error: RMSE) を用いた。値が小さいほど予測が正確であることを意味する。

が高いことを意味する。これをみると、モデル間の予測誤差には差がなく、ランダム・ウォーク・モデルの予測精度を超えるモデルはないことが窺われる²²。

本稿の実験のように、過去の株価に限定された情報集合の下で近年の日本の株式収益率の予測を行った場合には、標準的な NN が RW や ARMA と比較して予測精度の改善に寄与するという証拠は得られなかった²³。多くの先行研究では、日本の株式市場がウィーク型の意味で効率的であることを報告しており²⁴、本稿の結果はそれらと整合的な結果であると言える。NN のような柔軟なモデルを用いても、超過収益に結びつくような有益な株価予測は簡単ではないことが示唆される。

おわりに

資産価格の予測可能性は金融市場の情報効率性に依存するため、パターン認識能力に優れ、複数の分野で高い性能を発揮している NN を利用したとしても、予測精度の改善に結びつかないこと

があり得る。ウィーク型の意味で効率的である市場においては、本稿の実験のように、NN が予測精度の改善に寄与することは難しいと考えられる。これは、簡単な例で言い換えると、裏表が平等に出るコインを投げた結果を高精度で予測することが不可能であることと同様である。達成可能な精度の上限が理論的に定まっている問題に対していくら最適化を試みようと、破れない天井は存在するということである。

もっとも、発展途上国の株式市場や低流動性の資産、月単位や年単位といったより長期の収益率の予測においては、ウィーク型の市場効率性が成り立たないとの実証結果もある。こうした資産価格の予測に NN を適用すれば、超過収益を得られ

る可能性がある。

また、日本の株式市場がセミストロング型の意味で効率的かどうかを検証の余地がある。株価の過去の推移だけでなく、幅広い経済・金融指標やニュースのテキスト情報、衛星画像等の非構造データといった大きな情報集合を用いれば、NN が優れた予測を実現する可能性も否定できない。

いずれにせよ、NN の利点を発揮するには、金融市場の効率性の度合いを見極めることが重要となる。これを逆に言うと、従来の時系列分析に加えて、本稿で示したような機械学習の手法を発展・応用していくことで、市場の効率性の検証をより多角的に行うことが有用であると考えられる²⁵。

* 現金金融市場局

¹ この点は次の文献を参照。Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin, "Attention is All You Need," *Advances in Neural Information Processing Systems*, 2017, pp.5998-6008.

² NN は、理論的に任意の関数を任意の精度で近似できることが知られており、これは普遍性定理や万能近似能力と呼ばれている。

³ 機械学習の文脈では、正答率の理論的な限界はバイズ誤り確率 (Bayes Error Rate) と呼ばれる。この確率を解析的に求めることは難しく、目下得られている正答率が理論的な上限にぶつかっているかどうかを判断することは容易ではない。

⁴ 株価は非定常過程の一つである単位根過程に従うことが知られており、回帰モデルがよく当てはまると誤解してしまうことがある (見せかけの回帰)。このため、予測検証の際には株価の階差 (収益率) を予測できるかが焦点となる。

⁵ 数理モデルとしての NN は、生物のニューロン (神経細胞) のネットワークと明確に区別するため、人工ニューラル・ネットワーク (Artificial Neural Network: ANN) と呼ばれることもある。

⁶ NN におけるニューロンはパーセプトロン (Perceptron) と呼ばれる。

⁷ NN はその性質上、損失関数が全てのウエイトについて微分可能となりさえすれば、パーセプトロンを積み上げることで柔軟にネットワークを構築することができる。本稿で取り上げたもの以外の有名なネットワーク構造として、Residual Network や Self-attention Network 等がある。

⁸ 各ニューロンが下層のニューロンの出力値を重み付き平均する以外に、最大値を採用するといった処理も許容され、ノイズや画像の回転等の操作に対して頑健性が高い。

⁹ 誤差逆伝播法 (Back Propagation) という、ネットワークの出力値と実際のデータとの誤差の情報を、ネットワークの出力層から入力層へと順次伝えていく方法によって、各層のウエイトを最適化する。より具体的には、ネットワークの出力値とデータとの誤差を、ネットワークが持つ各ウエイトで偏微分し、誤差を小さくする方向にウエイトの更新を繰り返し、その収束先の値を学習済みのウエイトとする。この際、局所最適解への収束の阻止と計算の高速化を目的として、確率的勾配降下法 (Stochastic Gradient Descent) が一般に使用される。これは、データからランダムに抽

出したサンプルに基づいて計算した誤差を小さくする方法である。

¹⁰ この点は次の文献などを先駆に多くの研究で実証されている。Keown, Arthur J., and John M. Pinkerton, "Merger Announcements and Insider Trading Activity: An Empirical Investigation," *The Journal of Finance*, 36(4), 1981, pp.855-869.

¹¹ 効率的市場仮説については次の文献などを参照。Fama, Eugene F., "Efficient Capital Markets: a Review of Theory and Empirical Work," *The Journal of Finance*, 25(2), 1970, pp.383-417.

¹² 具体的には、価格の上昇や下降が連続する確率、反転する確率、上昇や下降の連続回数の統計検定 (Binary Run Length Test) や、分散比の検定 (Variance Ratio Test) である。

¹³ 具体的には、テクニカル分析や統計モデルに基づく予測のほか、予め定められた閾値よりも上昇 (下降) した際に買う (売る) というルールに基づいて売買するフィルタールールによる検証、ディーラーの売買ルールを模倣したエキスパートシステムによる検証、経済のファンダメンタル分析に基づく検証等がある。

¹⁴ 明日までの収益率よりも来年までの収益率の方が予測しやすいというのは一見不思議に思えるが、これは投資家が要求するリスクプレミアムが景気変動と密接に関係しており、従って株価収益率にも周期性が発生するためであると解釈がある。

¹⁵ 例えば、1月効果 (1月は他の月に比べて株価の平均的な収益率が高い)、曜日効果 (月曜日は他の曜日に比べて平均的な収益率が低い)、SAD (Seasonal Affective Disorder) 効果 (季節の変化に伴う日照時間の変化が収益率に影響) 等が指摘されている。

¹⁶ ランダム・ウォーク・モデルは、次式で表される。

$$p_{t+1} - p_t = \mu + \epsilon_t$$

ここで、 p_t は t 時点における株価の対数、 μ は合理的に期待される収益率、 ϵ_t は平均0の正規分布に従い、系列相関のないノイズである。合理的に期待される収益率とは、その資産のリスクに見合った収益率を指す。リスクを基に収益率を推計するモデルには、CAPM やファーマ・フレンチの3ファクター・モデル等、様々なものが提案されている。

¹⁷ ドリフト項には、モデルの推定に用いるデータ期間中の平均収益率を用いた。

¹⁸ ARMA(p, q) の p の最大値、 q の最大値をともに 10 とし、AIC に基づいて最適な次数を選択した。ここで p は株価収益率の自己

ラグの次数、 q は誤差項のラグ次数を表す。この結果、日経平均株価については ARMA(1,1)、日経平均先物については ARMA(0,1)、TOPIX については ARMA(3,1)、TOPIX 先物については ARMA(1,1) が最適なモデルとして推定された。

¹⁹ 同じ順伝播型 NN や畳み込み NN といっても、入力層のニューロンの数（予測に使う過去の収益率の日数）や中間層（入力層と出力層の間の層）のニューロンの数等、様々な設定があり得る。ここでは、3 種類の NN それぞれについて、入力層および中間層のニューロン数の候補を 2, 4, 8, 16, 32、層数（入力層と出力層を含む）の候補を 3, 4、用いる活性化関数の候補を Tanh（ハイパボリック・タンジェント）、ReLU (Rectified Linear Unit)、Sigmoid とし、データに対して最も当てはまりの良いものを選択した。ここで、Tanh は座標点を基点として点対称となる S 字型の曲線で、-1~1 の間の値を返す関数、ReLU は入力値 0 以下なら 0、0 超なら入力値と同じ値を返す関数、Sigmoid は座標軸を基点として点対称となる S 字型の曲線で、0~1 の間の値を返す関数である。当てはまりの良さの判断は、2000 年 2 月 3 日から 2017 年 12 月 27 日までのデータのうち、前 90% を用いて推定したモデルの、後 10% における予測精度に基づいて行った。この結果、どの種類の NN においても入力層のニューロン数は 2 が最適と推定された。

²⁰ ローリング推計と呼ばれる手法であり、金融時系列分析では一般的な方法である。

²¹ 予測精度に大きな差がなければ、超過的な収益は取引コストで相殺されてしまう可能性もある。

²² NN による予測誤差の方が RW よりも大きい傾向にある。この原因の 1 つとしては、NN が訓練データ（モデルの推定に用いるデータ）の特徴を過度に習得し、それ以外のデータへの当てはまりが低下する過学習 (Overfitting) が起こっている可能性が考えられる。しかし、予測誤差の違いはわずかであり、RW による予測に劣っているとまでは言えない。念のため、モデル間の予測誤差

に差があるかを検定 (Diebold-Mariano 検定) したところ、どの組み合わせについても統計学的に有意な差はみられなかった。

²³ 各株式収益率の時系列データに系列相関がないことは、統計検定によっても確かめられた。具体的には、Ljung-Box 検定を行ったところ、10 次までの自己相関係数がすべてゼロであるという帰無仮説を棄却できないことが確認された。

²⁴ 例えば Kim, Jae H., and Abul Shamsuddin, "Are Asian Stock Markets Efficient? Evidence from New Multiple Variance Ratio Tests," *The Journal of Empirical Finance*, 15(3), 2008, pp.518-532 では、1990 年以降の日本の株式市場について分散比検定の一種を実施し、ウィーク型の意味で効率的であると結論づけている。

²⁵ 機械学習を利用してセミストロング型の市場効率性を検証した研究として次の文献などがある。Gu, Shihao, Bryan Kelly, and Dacheng Xiu, "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, 33, 2020, pp.2223-2273.

日銀レビュー・シリーズは、最近の金融経済の話題を、金融経済に関心を有する幅広い読者層を対象として、平易かつ簡潔に解説するために、日本銀行が編集・発行しているものです。ただし、レポートで示された意見は執筆者に属し、必ずしも日本銀行の見解を示すものではありません。

内容に関するご質問等に関しましては、日本銀行金融研究所（代表 03-3279-1111）までお知らせ下さい。なお、日銀レビュー・シリーズおよび日本銀行ワーキングペーパー・シリーズは、<https://www.boj.or.jp> で入手できます。