

評価関数の可視化による株価予測モデルの汎用性評価

Versatility Evaluation of Stock Prediction Model by Visualization of Cost Function

坂下好希¹ 中井悠貴² 瀬之口潤輔³

Koki Sakashita¹, Yuki Nakai², and Junsuke Senoguchi³

¹ 東京大学大学院経済研究科

¹ Graduate School of Economics, The University of Tokyo

² 早稲田大学大学院先進理工学研究科

² School of Advanced Science and Engineering, Waseda University

³ 東京工科大学

³ School of Computer Science, Tokyo University of Technology

Abstract: When predicting stock prices with a complex model using machine learning or artificial intelligence, overfitting sometimes occurs, and the prediction accuracy expected in actual operation cannot be obtained. In such a model, the cost function is presumed to be steep and multi-modal, while in a model that maintains stable prediction results, the cost function is considered to be gradual and single-peaked. In this study, we first compared the performance of several stock price prediction models, and then visualized the cost function for each model using t-SNE. As a result, the model using Lasso regression, which had the highest performance, showed a gradual unimodal cost function, while the linear regression, which had relatively low performance, showed a steep and multi-modal shape.

Visualizing the cost function using t-SNE can be an important index for evaluating the stability and versatility of a stock price prediction model.

1. はじめに

株価予測は多くの研究者や実務家によって行われている研究領域である。近年では伝統的な統計手法に加えて機械学習や人工知能を用いて、複雑なモデルを構築し株価予測を行う研究も多い。実務でも、人工知能により株価判断を行い、資金を運用する投資信託が登場している。しかし機械学習等を用いた複雑なモデルにより株価が予測可能かについては、依然として結論は出ていない。株価予測モデルに頻繁にみられる問題として、訓練時では高い予測精度を示す一方で、実際に運用では期待されていた予測精度が得られないことである。これには複雑なモデルを構築することにより訓練データのノイズ等まで学習してしまうオーバーフィッティングが要因となっていることがある。

特に株価のように様々な非連続的な要因により変動する対象をモデル化する場合、パラメータの数を増やし複雑なモデルで訓練データを学習すると、最適なパラメータによる学習データの評価値は高いも

の、非連続的なノイズまで学習している可能性から、その周辺のパラメータを用いた評価値は大幅に悪くなることもある[1]。パラメータの最適値からわずかに変化するだけで、評価値が大きく悪化するということは、最適モデルにおいて訓練データがわずかに変化するだけで評価値が大きく悪化することにもなる。つまりモデルに汎用性が乏しいことを意味する。このようなモデルを用いて将来を予測しても、十分に高い予測精度を示すことができないと考えられる(図1a)。

一方で最適なパラメータから多少離れても評価値が大きく変わらない、つまり評価関数のグラフが比較的なだらかなモデルであった場合、訓練データが多少変わっても評価値は大きく変化せず、モデルに汎用性が高いことを意味する。このようなモデルを用いて将来を予測した場合、比較的高い予測精度を示すことができると考えられる(図1b)。

このように評価関数を可視化してモデルの構造を解明することは、モデルの安定性・汎用性を評価するうえで極めて重要である。しかしパラメータが高

次元である場合、評価関数も非常に複雑な形状となり、人間がそれを直接的に理解することは極めて困難である。よって本研究では、高次元空間におけるデータ間の関係性やデータの特徴を保持したまま、低次元空間にマッピングする手法である t-SNE (t-Distributed Stochastic Neighbor Embedding)を用いて、高次元モデルの評価関数を可視化し、その形状がなだらかな単峰性であればモデルの安定性・汎用性が高く、結果としてある一定期間だけではなく長期間にわたり運用成績が安定する株価予測モデルを作成することができる。

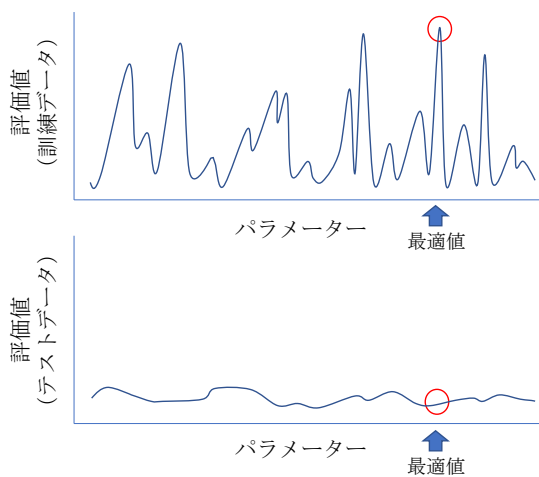


図 1a : 評価関数の形状とモデルの汎用性：急峻な多峰性の場合

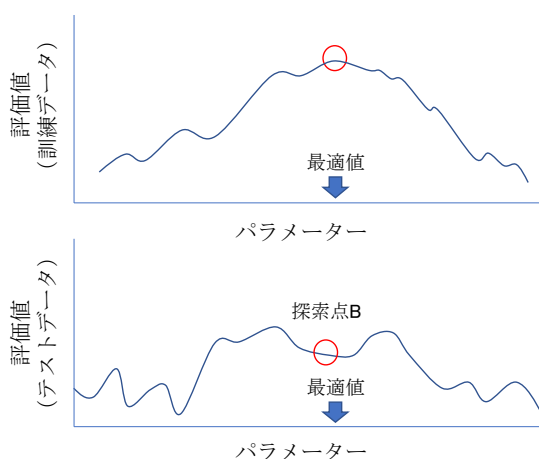


図 1b : 評価関数の形状とモデルの汎用性：緩やかな単峰性の場合

以上のことから、本研究の第一の目的は、ある一定期間だけではなく長期間にわたり運用成績が安定する株価予測モデルを作成することとする。また第

二の目的は、株価予測モデルの評価関数の構造を可視化し、緩やかな単峰性のある評価関数をもつ株価予測モデルが安定性・汎用性および運用成績において優れた評価を得ることができるか検証することである。

2. 先行研究

株価予測モデルの精度を証明するためには、モデルの構造を解明してモデルの汎用性が高いことを確かめることが重要である。モデル構造の解明には、高次元データを次元圧縮し可視化することが大いに役立つ。

高次元データの次元圧縮とその可視化は、機械学習やデータマイニングの分野において古くから非常に重要な問題であり、これまでに多くの手法が開発されてきた。次元圧縮の代表的な手法である主成分分析は[2]により導入され、可視化手法として[3]による Multidimensional scaling が開発された。

高次元空間上でデータが同一の非線形多様体上に存在する場合、線形の次元圧縮手法ではマッピング先の低次元区間上において、それらのデータを類似しているものとして表現することができない。このため 2000 年代初頭には非線形の次元圧縮手法が次々に開発され、代表的なものとして[4]の IsoMap, [5]の Locally Linear Embedding, [6]の Laplacian Eigenmaps などがある。

しかし多くの非線形次元圧縮手法は人工データには有効であるが、実データに対しては精度が著しく低下するという欠点があった。[7]はこの問題を解決する手法として、t-SNE の開発を行った。t-SNE は、人工データだけでなく複雑な実データに対しても有効であり、かつより高次元のデータに対しても適切に機能することが示されている。

3. データと分析手法

3.1 分析に用いたデータ

目的変数として、期近の TOPIX 先物の翌日始値 (8:45) から終値 (15:15) のリターンを用いる。

説明変数として、表 1 に示した金融指標を用いた。

変数 No.1~No.7 からウェーブレット変換により各レベルのウェーブレット係数を抽出する。ウェーブレット係数のレベルは 1~3 とする。つまりサポート期間が 2 営業日, 4 営業日, 8 営業日とする。ウェーブレットフィルタはハールウェーブレットを使用する。それらに 4 種類のテクニカル指標を加えた 25 種類の指標を説明変数として採用する。

表 1：分析に使用する金融指数と項目番号

項目番号	項目名	基準	単位
No.1	TOPIX (東証株価指数)	前日東京終値～当日東京終値	変化率
No.2	S&P 500種	前日NY終値～当日NY終値	変化率
No.3	S&P 500コース指数	前日NY終値～当日NY終値	対S&P相対変化率
No.4	S&P 中型株400種	前日NY終値～当日NY終値	対S&P相対変化率
No.5	CBOE SPXボラティリティ指数	前日NY終値～当日NY終値	変化幅
No.6	米ドル/円	前日NY終値～当日NY終値	変化率
No.7	TOPIX過去20日移動平均からの乖離	当日東京終値	水準
No.8	TOPIX14日RSI	当日東京終値	水準
No.9	14日RSIが70を上回っているTOPIX構成銘柄の割合	当日東京終値	水準
No.10	14日RSIが30を下回っているTOPIX構成銘柄の割合	当日東京終値	水準

3.2 モデルの予測精度の評価方法

本研究では、モデルの予測精度の評価指標として、walk forward test による累積リターンを採用する。

walk forward test では、予測を行う当日まで T 日分の訓練データを用いてモデルを学習して、翌日の予測値を算出する。予測は翌日分だけ行う。1 日ずらして同様の操作を中長期にわたり繰り返していく。この手法を採用した理由は、毎日モデルを作り直していくので、株式市場の局面が変化したとしてもモデルがそれに対応するため、モデルが陳腐化しにくいことである。また回帰モデルの回帰係数の時間変化を確認できるため、株価予測モデルの構造変化を確認できるメリットもある。

T に関しては、[8]が 2500 日分の株価データ、[9]が約 1600 日分のデータでそれぞれモデルを学習させていたことを参考に、この 2 つの中間値に近い 2000 日とした。

説明変数と目的変数はそれぞれ標準化してモデルを学習させ、翌日の標準化された予測値を計算する。得られた予測値に標準化の逆の操作を行い、これを翌日のリターン予測値とする。モデルは Lasso 回帰 ($\alpha=0.01$)、Ridge 回帰 ($\alpha=10$)、Elastic Net 回帰 ($\alpha=0.01, p=0.5$)、正則化項を用いない重回帰分析 (Linear 回帰) の 4 種類を用いる。

4 種類それぞれのモデルで、予測精度の評価を行う。具体的には、翌日のリターン予測値が正であれば TOPIX 先物を翌日の寄付き (始値) で買い、引け (終値) で売る。負であれば翌日の寄付き (始値) で売り、引け (終値) で買い戻す。これを walk forward test の全期間で繰り返す。walk forward test により翌日のリターン予測を行う期間は 2001 年 9 月 10 日から 2019 年 2 月 13 日までとする。つまり 4,274 回の予想を行い、その成績に基づきモデルを評価する。

2001 年 9 月 10 日運用前の資産を 1 として、上記

のルールに基づいて複利で運用する。2019 年 2 月 13 日時点の資産を株価予想モデルの累積リターンとし、これを株価予測モデルの予測精度の評価基準とする。

評価の基準は、(1) 4,274 回の騰落予測の正解率 (acc)、(2) 4,274 回のうち翌日のリターンの絶対値が 2%以上のとき騰落予測の正解率 (acc_2%)、(3) 4,274 個の株価予想と実績に関する平均平方二乗誤差 (RMSE)、(4) 2001 年 9 月 10 日運用前の資産を 1 として、上記のルールに基づいて複利で運用した場合の累積リターン (cum_r) の 4 つとする。

3.3 モデルの安定性の評価手法

本研究では、予測精度の高い株価予測モデルを作成するだけでなく、株価予測モデルの安定性・汎用性の評価も行う。

本研究で用いるモデルは翌日の株価騰落を予測するものであり、短期的な時系列データ予測を検証しているといえる。しかしながら、株価予測においては株価のランダムネスや突発的な変動により、必ずしもある一定期間の短期的な予測精度が中長期的な累積リターンにつながるわけではないということが明らかとなっている。そこで我々は、本研究でモデルの精度の評価基準となっている中長期的な累積リターンに注目して、モデルの安定性・汎用性を検証する。

モデルの安定性・汎用性を評価する手法を図 3 に示す。本研究では walk forward test により、一日ごとに翌日の株価を予測するためのモデルを獲得することができる (これを「予測モデル」と呼ぶ)。そして、その 1 日ごとの予測モデルを用いて、n 日間運用した時に得られるモデルを「運用モデル」とする。このとき、予測モデル学習における説明変数の数を m とすると、運用モデルのパラメータは $n \times m$ 個のマトリックスとして与えられる。このようにして得ら

れた運用モデルの $n \times m$ 個のパラメータにそれぞれ異なるガウシアンノイズ (標準偏差=0~0.1) を与え、ノイズ付きの運用モデルによる累積リターンにより、モデルの安定性・汎用性を評価する。つまりモデルの費用関数が最適値付近でなだらかな単峰性の特徴を有していれば、パラメータが多少変化しても累積リターンはさほど変化しない。このような安定した費用関数を持つモデルは、市場環境が変わったとしても安定した予想精度を維持することができ、汎用性が高いと言える。一方で費用関数が急峻で多峰性なものである場合、パラメータの値が最適値から多少ずれるだけで累積リターンは急激に悪化する。このような費用関数を持つモデルは汎用性が低いと言える。これらのことから、運用モデルのパラメータにノイズを加えたときの累積リターンの大きさと

変動係数を、モデルの安定性・汎用性を評価する指標として用いる。

更に、t-SNE を用いて運用モデルを 2 次元空間に圧縮することによって、運用モデルと累積リターンとの関係性について調べる。本評価では、異なるノイズを付与した運用モデル ($n \times m$) を k 個作成する、そして、 k 個のモデルに対して t-SNE を施すことにより 2 次元空間に描画する。つまり、 k 個の運用モデル $M(M_i \in R^{n \times m}, i = 1, 2, \dots, k)$ を 2 次元空間 (R^2) にマッピングするのである。更に、それぞれの運用モデルによる累積リターンと t-SNE によって描画された 2 次元空間での距離から、累積リターンと運用モデルの関係性について議論する。

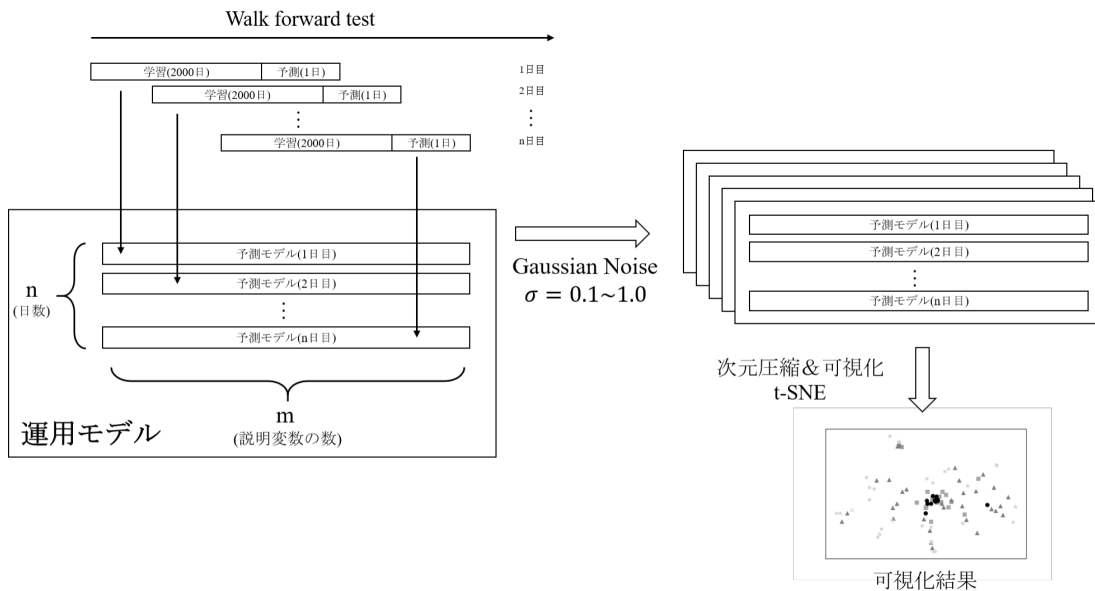


図 2 : モデルの安定性の評価方法

4. 分析結果

4.1 モデルの予測精度に関する評価結果

表 2 に各モデルの予測精度の一覧を示す。予測精度の評価方法は、4.2 で示した 4 つの基準を用いた。

Lasso 回帰がいずれの指標においても、最も高いパフォーマンスとなっている。次いで、Elastic Net 回帰のパフォーマンスが高く、Ridge 回帰と正則化項を用いない重回帰分析 (Linear 回帰) はいずれの指標においてもパフォーマンスが劣っていた。正則化項を用いることで精度が高まり、特に L1 正則化項による次元圧縮がモデルのパフォーマンスを改善さ

せていることが確認された。

表 2 : 各モデルの精度

	acc	acc_2%	RMSE	cum_r
Lasso	51.8%	58.4%	0.01022	40.1
Ridge	51.1%	55.9%	0.01026	17.6
Elastic Net	51.2%	57.7%	0.01023	25.7
Linear	51.1%	55.5%	0.01026	16.5

図 3 に各モデルで walk forward test を行ったときの累積リターンの推移を示す。ベンチマークとして TOPIX の推移も示した。いずれのモデルでも TOPIX

をアウトパフォームしており、中でも Lasso 回帰のモデルは約 20 年間の walk forward test 期間で資産価値は約 40 倍となった。騰落予測の正解率が 51.8%と高くない水準で累積リターンが大きい理由は、acc_2%が 58.4%と高いためである。つまり本研究で作成した株価予測モデルは、安定性・汎用性を重視しているため、一過性の要因による株価の微小な変化を正確に予測することはできないが、一方で株価の大きな変化の要因になるような普遍的なコアの構造は概ね捉えており、これが walk forward test 全体の正解率が高くない一方で acc_2%が比較的高い理由となっている。

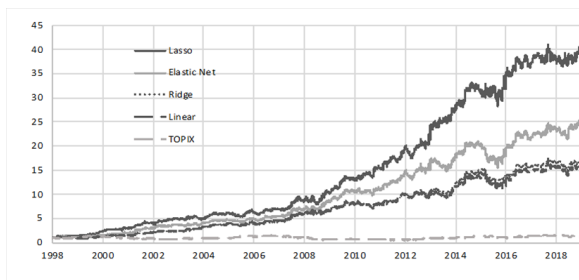


図 3：各モデルの累積リターンの推移

4.2 ノイズ付与によるモデルの安定性・汎用性の評価結果

図 4 は、運用モデルに与えるガウシアンノイズの標準偏差ごと (0~0.1) に、累積リターンを比較したものである。それぞれの標準偏差で異なるノイズの付与を 10 回行い、それから得られる 10 個の累積リターンの平均値を示している。Ridge 回帰と Linear 回帰は、いずれの標準偏差においても累積リターンは低い水準である。一方で Lasso 回帰と Elastic Net 回帰は、累積リターンの水準が比較的高く、またノイズ大きくなるに伴い累積リターンが低下していることがわかる。

図 5 は、図 4 の結果のもとになる 10 個の累積リターンの変動係数を示したものである。Lasso 回帰を用いた分析の変動係数が安定して低く推移していることがわかる。Elastic 回帰を用いた分析では、累積リターンは大きいですが、その安定性は低いことが示された。また、Ridge 回帰と Linear を用いた分析では、累積リターンの水準が低いことに加えて変動係数も高く、予測精度および安定性ともに低いことが示された。

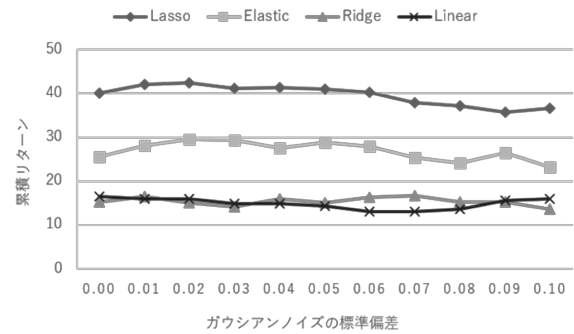


図 4：ノイズ付与による累積リターンの変化

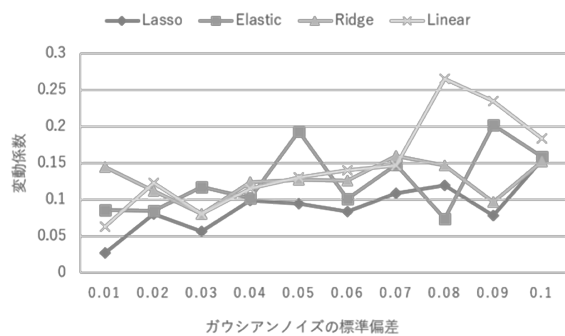


図 5：ノイズ付与時の累積リターンの変動係数

4.3 t-SNE を用いたモデルの安定性・汎用性の評価結果

図 6 は、ガウシアンノイズ (標準偏差=1) を加えた運用モデルを t-SNE を用いて 2 次元空間上に次元圧縮し、散布図として表したものである ((a) Lasso 回帰, (b) Elastic Net 回帰, (c) Ridge 回帰, (d) Linear 回帰)。次元圧縮の際の用いたパラメータはパープレキシティを 50, イテレーションを 1000 に設定した。各プロットはガウシアンノイズを付与した運用モデルを表しており、その総数は 100 である。また、プロットの形状はそれぞれの運用モデルを用いた際の累積リターンを表している (円: 40 倍以上, 四角形: 30~40 倍, 三角形: 20~30 倍, 星形: 10~20 倍, ×印: 10 倍以下)。

図 6(a)では、中心部に累積リターンが 30 倍以上の運用モデルが集まっており、中心から遠ざかっていくにつれて、累積リターンは下がっていることがわかる。つまり、Lasso 回帰を用いたモデルでは、高次元空間におけるモデルの類似性と累積リターンの大きさには相関があるといえる。一方、図 6(b), (c) と(d)では、次元圧縮後の各モデルの位置と累積リターンの間に明確な規則性は見られなかった。

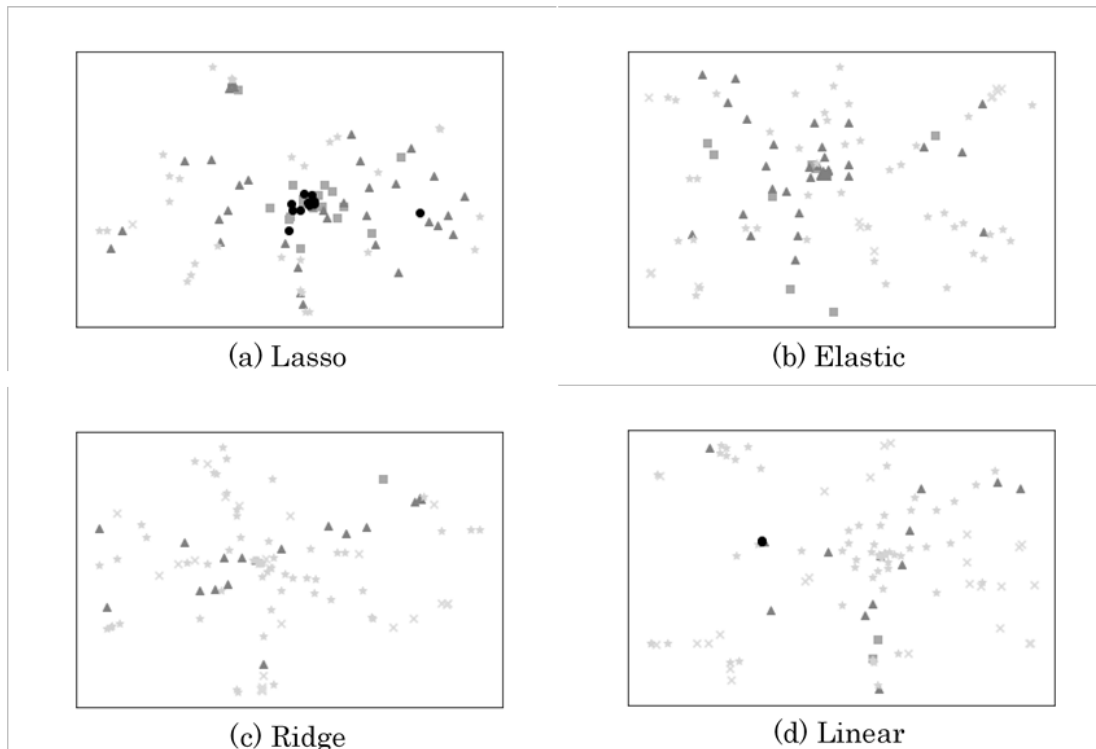


図 6 : t-SME による可視化結果

5. 考察

本研究の目的は、長期間にわたり運用成績が安定している株価予測モデルを作成することに加えて、その構造を可視化し解明することにより、どのようなモデルが安定した運用成績を継続できるか、評価の指針を示すことである。

まず運用成績が安定している株価予想モデルの作成について、表2および図3から、正則化項を用いて次元圧縮を行うことが効果的であることが示された。株式市場では多数の参加者が価格形成に影響を与えるため、株価は短期間で乱高下することがあり、また金融環境は常に変化しているためこのような短期間の動きに影響を与える要因も一定とは限らない。このような株式市場の特性を多数のパラメータを用いた複雑なモデルで表現した場合、過去の一過性のノイズまで説明することによりモデルが過学習を起しやすいため複製ができず、データ数の制約がある中で過学習の問題を回避するには、パラメータを重要なものに制限し株式市場のコアの構造を捉えるモデルを構築することが特に重要である。正則化項を用いて次元圧縮を行ったモデルの予測精度が高い結果となったのは、モデルがノイズを拾わずに、株式市

場の普遍的なパターンを表現しているためと考えられる。

次に安定した運用成績を継続できるかについて、構築したモデルのすべてのパラメータに対してノイズを付与し運用成績がどの程度変化するかを計測することにより、モデルの汎用性と安定性を評価した。図4の結果から、Lasso 回帰を用いた手法はノイズの大きさに依存して累積リターンが変動しているものの、基本的に他の手法よりも高い累積リターンを得られることが分かった。また、累積リターンの大きさは、明らかに Lasso 回帰 > Elastic Net > Ridge 回帰 \approx Linear 回帰となっている。本研究で用いた説明変数数は 24 であり、この程度の説明変数数では L_1 正則化項を用いた学習が適していることが示された。

更に図5の結果から、Lasso 回帰を用いた手法は他の正則化を用いた手法に比較して、モデルに同じ標準偏差のノイズを加えた場合の累積リターンの変化が小さく、図4と同様にパラメータに付与したノイズに対して高い安定性を示している。これは株式市場にノイズが発生しても、予測結果が影響を受けにくいことを示しており、長期間にわたり安定した運用成績を継続できる要因と考えられる。Lasso 回帰は、必要のない説明変数を 0 とするためにスパース性が高く、これにより他の次元圧縮手法に比べてもノイ

ズの影響を受けにくい構造となっていると考えられる。

Lasso 回帰は予測精度、安定性ともに高く、また Linear 回帰はともに低い。これを図 6 の結果と照らし合わせると、費用関数の形状が予測精度と安定性に影響を与えていることがわかる。Lasso 回帰では高次元の空間における運用モデルの類似度と累積リターンの大きさの間には依存関係があることがわかる。つまり、Lasso 回帰による株価予測モデルの費用関数は、単峰性のなだらかな形状をしていることがわかる。これは図 1(a)で示した安定したモデルの構造を反映したものであり、それゆえ株価予測の精度も高く、モデルの安定性も比較的高い結果となっていると言える。これに対して、Linear と Elastic, Ridge による手法では、凹凸が入り組んだような複雑なモデルが得られたことがわかる。つまり費用関数は各所に評価値が高い急峻な山があり、図 1(b)で示したような急峻で多峰性の評価関数になっている。このような費用関数を持つモデルは訓練データに対して過学習を起こしていると考えられ、それゆえ株価予測の精度が低く、モデルの安定性も低い結果になっていると言える。

このようにモデルのすべてのパラメータにノイズを付与し、t-SNE により 2 次元にマッピングすることにより、高次元モデルの費用関数を可視化できる。費用関数の形状がなだらかな単峰性である場合は、長期間にわたり運用成績が安定している株価予想モデルと評価することができ、また急峻な多峰性である場合は、安定性、運用成績ともに不確実な株価予想モデルと評価することができる。

6. おわりに

機械学習や人工知能を用いて複雑なモデルを構築し株価予測を行う場合、訓練データのノイズ等まで学習してしまうオーバーフィッティングが要因となり、訓練時では高い予測精度を示す一方で、実際の運用では期待されていた予測精度が得られないことがある。

オーバーフィッティングを起こしている株価予測モデルでは、評価関数が急峻な多峰性であると考えられ、また長期間にわたり安定した運用成績を継続している株価予測モデルでは、評価関数がなだらかな単峰性であると考えられる。

本研究では、まず複数の株価予測モデルにより長期間にわたる運用成績を比較し、それぞれのモデルについて評価関数を t-SNE を用いて可視化した。この結果、運用成績が最も高かった Lasso 回帰では、評

価関数がなだらかな単峰性を示しており、運用成績が比較的良かった Ridge 回帰または Linear 回帰では評価関数が多峰性であった。このことから、t-SNE を用いて評価関数を可視化することは、株価予測モデルの安定性や汎用性を評価する指標になり得ることが示された。

本研究では評価関数の形状を目視により判断したが、関数の多峰性を検定する手法は存在する。評価関数の形状を多峰性の検定を用いて定量的に評価し、モデルの安定性と予測精度をより正確に示すことができる指標を作成することを、今後の研究の課題とする。

参考文献：

- [1] 松井和宏・佐藤晴夫「GA による株取引戦略獲得における近傍評価法の改良」『人工知能学会全国大会』第 24 回, 2010 年.
- [2] Pearson, K.: “On Lines and Planes of Closest Fit to Systems of Points in Space,” *Philosophical Magazine*, Vol. 2, No. 11, pp. 559–572, (1901)
- [3] Torgerson, W. S.: “Multidimensional scaling: I. theory and method,” *Psychometrika*, Vol. 17, No. 4, pp. 401–419, (1952)
- [4] Tenenbaum, J. B., De Silva, V., and Langford, J. C.: “A Global Geometric Framework for Nonlinear Dimensionality Reduction,” *Science*, Vol. 290, No. 5500, pp. 2319–2323, (2000)
- [5] Roweis, S. T., and Saul, L. K.: “Nonlinear Dimensionality Reduction by Locally Linear Embedding,” *Science*, Vol. 290, No. 5500, pp. 2323–2326, (2000)
- [6] Belkin, M., and Niyogi, P.: “Laplacian Eigenmaps and Spectral Techniques for Embedding and Clustering,” In *NIPS*, Vol. 14, pp. 585–591, (2001)
- [7] Maaten, L. v. d., and Hinton, G.: “Visualizing Data using t-SNE,” *Journal of Machine Learning Research*, Vol. 9 (Nov), pp. 2579–2605, (2008)
- [8] Roy, S. S., Mittal, D., Basu, A., and Abraham, A.: “Stock Market Forecasting Using LASSO Linear Regression Model,” *Afro-European Conference for Industrial Advancement, Advances in Intelligent Systems and Computing*, Vol. 334, pp. 371–381, (2015)
- [9] 松本健・牧本直樹「LSTM による時系列予測と株式投資戦略への応用」『人工知能学会金融情報研究会』第 22 回, 2019 年.