

# 正則化最小二乗法を用いた 多変量多項式回帰モデルに対するパラメータ推定法

井上 一磨<sup>1,2,a)</sup> 清水 良太郎<sup>3</sup> 須子 統太<sup>4</sup> 後藤 正幸<sup>3</sup>

受付日 2023年5月31日, 再受付日 2023年7月21日/2023年8月28日,  
採録日 2023年9月15日

**概要:** 多項式回帰モデルは、単純な線形回帰モデルに比べて複雑な関係を表現できるモデルであり、様々な予測問題に活用されている。また、複雑な処理を行う機械学習モデルに比べて、高い解釈性を持つという点でも有用なモデルであると考えられる。しかし、多変量の多項式回帰モデルでは、次数が大きくなると変数の数が指数的に増大するため、適切な説明変数の選択が必要である。従来、多変量多項式回帰モデルに対してスパース推定を用いて、効率的に高次の項の推定を可能にするアルゴリズムが提案されている。それらの1つである SPORE-LASSO は、元の説明変数に LASSO を適用して一部の変数を取り出し、少数の説明変数に任意の次数の多項式展開を行って構築した多項式回帰モデルに対して、再度スパース推定を行うアルゴリズムである。その結果、すべての説明変数を考慮せず、多項式回帰モデルに対する効率的なパラメータ推定を実現している。しかし、従来のアルゴリズムは、最大次数を固定する必要がある。故に、最大次数を小さく設定すると真の構造をとらえることができなくなってしまい、大きく設定した場合には上述のように、変数の数が指数的に増加してしまうという問題を抱えている。そこで本稿では、最大次数が未知の多項式回帰モデルに対し、スパース推定を行う新たなアルゴリズムを提案し、人工データによるシミュレーション実験と実データを用いた実験によりその有効性を検証する。

キーワード: 予測アルゴリズム, 多項式回帰モデル, LASSO

## An Efficient Approach to Parameter Estimation in Multivariate Polynomial Regression Models via Regularized Least Squares

KAZUMA INOUE<sup>1,2,a)</sup> RYOTARO SHIMIZU<sup>3</sup> TOTA SUKO<sup>4</sup> MASAYUKI GOTO<sup>3</sup>

Received: May 31, 2023, Revised: July 21, 2023/August 28, 2023,  
Accepted: September 15, 2023

**Abstract:** We introduce a novel algorithm designed to estimate the parameters of multivariate polynomial regression models with an undefined maximum order. Polynomial regression models are crucial tools in predictive tasks. However, as the number of explanatory variables rises, so does the complexity of the polynomial terms, making accurate parameter estimation particularly challenging when dealing with higher order models. The SPORE-LASSO algorithm tackles this issue by leveraging sparse estimation for multivariate polynomial regression models, effectively approximating higher order terms. By utilizing the LASSO technique, it identifies a concise set of significant explanatory variables. This method permits sparse estimation of polynomial regression models, adeptly managing a restricted number of polynomial terms. Nonetheless, conventional algorithms mandate the specification of a maximum order. When this parameter is set too low, the model's true structure may be overlooked. Conversely, if the maximum order is set too high, an unwieldy increase in the number of variables arises. Thus, we've developed a new algorithm that mitigates these issues. We have confirmed the efficacy of this algorithm through simulation experiments employing both synthetic and real-world data.

**Keywords:** prediction algorithm, polynomial regression model, LASSO

## 1. はじめに

多項式回帰モデル [1] は、単純な線形回帰モデルと比べ、複雑な関係を表現することができるモデルであり、様々な予測問題に対する適用事例が報告されている [2], [3]. 複雑な機械学習モデルに比べて構造が単純であり、モデルの解釈性 [4] に富んでいる。近年、モデルが実ビジネスにおいて運用されるにつれ、機械学習モデルからの出力や予測値に基づいた意思決定が行われるようになった。しかし、多くの機械学習モデルは複雑な処理を行っており、内部の仕組みを理解することが難しく、解釈性の低さが運用上の問題としてあげられている [5], [6]. たとえば、金融業界の信用リスク判定モデルでは、個人の信用評価（信用できるか、否か）に対して、なぜその判断を下したかに関する要因を示すことが求められる。それゆえ、解釈性の欠如は、モデルの信頼性や公平性に影響を与える可能性があることから、解釈性のある多項式回帰モデルは非常に有用なモデルであると考えられる。しかし、多変数の多項式回帰モデルでは、次数が大きくなると変数の数が指数的に増大するため、適切な変数選択を行う必要がある。

従来、多変数の多項式回帰モデルに対してスパース推定を用いることで、効率的に高次項を含むモデルのパラメータ推定を可能にするアルゴリズムが提案されている。また、実データに対して有効な事例が多数報告されている。Rendle [7] は、高次元でスパースなデータに対して交互作用項を含んだ新しいモデルクラスとして Factorization Machines を提案した。このアルゴリズムでは、交互作用項のパラメータを低次元で表したベクトルの内積で表現することで、パラメータ推定の効率化を実現している。また、Huang ら [8] は、SPORE-LASSO というアルゴリズムを提案した。この方法ではまず、入力した説明変数から LASSO [9] によって少数の変数を選択する。次に、一部の説明変数に対して任意の次数で多項式展開を行い、多項式回帰モデルを構築する。最後に、再度スパース推定を行うことで、すべての変数を用いて多項式回帰モデルのパラメータ推定を行う必要がなくなり、効率的に多変数多項式回帰モデルの構築が可能になった。SPORE-LASSO を、コンピュータプログラムの実行時間を予測する問題に対して適用し、効率性だけでなく、予測精度の観点でも比較手法

よりも高い予測精度を示すことが報告されている [8]. さらに、Bertsimas ら [10] も同様のアプローチを採用したアルゴリズムを提案しており、一部の説明変数を抽出して多項式回帰モデルを構築したのち、LASSO のパラメータ推定時に整数最適化問題を解いている。この最適化方法では、変数間の階層的な構造を考慮できることから、LASSO に比べ、よりスパース性の高いパラメータ推定が可能となる。

上記のように、従来提案されている多項式回帰モデルに対するスパース推定は、最大次数をあらかじめ設定する必要がある。そのため、真の次数が事前に設定した次数よりも大きい場合には、真の構造をとらえることができないという問題を抱えている。また、最大次数を大きくし過ぎると、計算量が非常に大きくなってしまいう問題がある。

そこで本稿では、最大次数が未知の場合において多項式回帰モデルに対し、効率的にスパース推定を行うアルゴリズムを構成し、その精度について検証を行う。また、人工データによるシミュレーション実験と実データを用いた実験双方により、パラメータの推定精度と予測精度に関して有効性の検証を行い、提案アルゴリズムの挙動に関する考察を与える。

## 2. 多変数回帰モデルに対するスパース推定

### 2.1 多変数多項式回帰モデル

本稿では、サイズが  $N$  のサンプル  $\{\mathbf{x}_i, y_i\}_{i=1}^N$  を観測している状況で、新たに取得した  $\mathbf{x}_{N+1}$  を用いて  $\hat{y}_{N+1}$  を予測する問題を考える。 $y_i \in \mathbb{R}$  は  $i$  番目の目的変数を表す。また、 $P$  個の説明変数を持つ  $i$  番目の説明変数ベクトルを  $\mathbf{x}_i = (1, x_{i1}, \dots, x_{iP})^\top \in \mathbb{R}^{P+1}$  と表す。

このとき、線形回帰モデルは以下のように表される。

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \dots, N. \quad (1)$$

ここで、 $\boldsymbol{\beta} \in \mathbb{R}^{P+1}$  は  $P+1$  次元のパラメータベクトルであり、観測誤差  $\varepsilon_1, \dots, \varepsilon_N$  は互いに独立に平均 0、分散  $\sigma_\varepsilon^2$  の正規分布に従うものとする。

データの生成過程を式 (1) と仮定した場合、パラメータの推定値  $\hat{\boldsymbol{\beta}} \in \mathbb{R}^{P+1}$  を得た後、新規データ  $\mathbf{x}_{N+1} \in \mathbb{R}^{P+1}$  に対して、下記のように予測値を得ることができる。

$$\hat{y}_{N+1} = \mathbf{x}_{N+1}^\top \hat{\boldsymbol{\beta}}. \quad (2)$$

パラメータの推定値を得る方法としては、最小二乗法等が考えられる。ここで、 $N$  個の目的変数を結合し、目的変数ベクトル  $\mathbf{y} = (y_1, y_2, \dots, y_N)^\top \in \mathbb{R}^N$  を構成し、さらに  $N$  個の説明変数ベクトルを並べることで計画行列  $\mathbf{X} \in \mathbb{R}^{N \times (P+1)}$  を得る。

<sup>1</sup> 早稲田大学社会科学部  
Graduate School of Social Sciences, Waseda University, Shinjuku, Tokyo 169-8050, Japan

<sup>2</sup> 株式会社エクサウィザーズ  
ExaWizards Inc., Minato, Tokyo 105-0021, Japan

<sup>3</sup> 早稲田大学創造理工学研究所  
Graduate School of Creative Science and Engineering, Waseda University, Shinjuku, Tokyo 169-8555, Japan

<sup>4</sup> 早稲田大学社会科学総合学術院  
Faculty of Social Sciences, Waseda University, Shinjuku, Tokyo 169-8050, Japan

a) lotte01433@fuji.waseda.jp

$$X = \begin{bmatrix} \mathbf{x}_1^\top \\ \mathbf{x}_2^\top \\ \vdots \\ \mathbf{x}_N^\top \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1P} \\ 1 & x_{21} & \cdots & x_{2P} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \cdots & x_{NP} \end{bmatrix}. \quad (3)$$

最小二乗推定量  $\hat{\beta}_{LS}$  は、次式で定義される。

$$\hat{\beta}_{LS} = \arg \min_{\beta} \frac{1}{2} \|\mathbf{y} - X\beta\|_2^2. \quad (4)$$

上記の最小化問題を解くことで、以下のように最小二乗推定量  $\hat{\beta}_{LS}$  が求められる。

$$\hat{\beta}_{LS} = (X^\top X)^{-1} X^\top \mathbf{y}. \quad (5)$$

次に、Tibshirani [9] によって提案された LASSO 推定量  $\hat{\beta}_{LASSO}$  は次式で定義される。なお、 $\beta_j$  はパラメータ  $\beta$  の第  $j$  成分を指す。

$$\hat{\beta}_{LASSO} = \arg \min_{\beta} \left( \frac{1}{2} \|\mathbf{y} - X\beta\|_2^2 + \lambda \sum_{j=1}^{P+1} |\beta_j| \right). \quad (6)$$

LASSO では、正則化項としてパラメータの絶対値和を損失関数に加えることで、目的変数に対して影響が小さいパラメータを 0 と推定する。その結果として、パラメータの推定と変数選択を同時に行うことができ、サンプルサイズよりも変数の数の方が多く場合でもパラメータの推定が可能となる。ただし、0 と推定されるパラメータの数は設定するハイパーパラメータ  $\lambda$  の値に依存し、 $\lambda$  の値を大きくしていくと、0 と推定されるパラメータの数が増えていく。

上記、単純な線形回帰モデルに対し、多変量多項式回帰モデルは、複数の説明変数と目的変数の間に存在する関係を多次元の多項式関数で表すことで、複雑な事象を表現することが可能である。ただし、多項式回帰モデルのパラメータを推定するには、元の計画行列  $X$  に対して多項式展開を施し、高次の項や交互作用項を含んだ新たな説明変数を得る必要がある。たとえば、最大次数  $d = 2$ 、変数の数  $P = 2$  の新たな計画行列  $X_{new} \in \mathbb{R}^{N \times \binom{P+d}{d}}$  とおくと、

$$X_{new} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{11}x_{12} & x_{11}^2 & x_{12}^2 \\ 1 & x_{21} & x_{22} & x_{21}x_{22} & x_{21}^2 & x_{22}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N1}x_{N2} & x_{N1}^2 & x_{N2}^2 \end{bmatrix}, \quad (7)$$

となる。上記のように多項式展開して得られた  $X_{new}$  を用いて式 (4)–(6) の行列部分を置換し、パラメータの推定量を得ることができる。このように技術的には多項式回帰モデルのパラメータの推定は可能である。なお、パラメータベクトル  $\beta$  の次元は、変数の数  $P$  と次数  $d$  を用いて  $\binom{P+d}{d}$  次元のベクトルで表すことができる。上記の

関係で示したように説明変数の数  $P$  と次数  $d$  が大きくなるに従って推定する項の数が増加してしまう。その結果、非常に大きいデータが対象である場合、計算量が膨大になってしまうといった問題を抱えている。

## 2.2 SPORE-LASSO

従来、上記であげた問題に対して効率的に多項式回帰モデルの推定を行うアルゴリズムとしてあげられるのが Huang ら [8] によって提案された SPORE-LASSO である。以下に、SPORE-LASSO のアルゴリズムを示す。

---

### Algorithm 1 SPORE-LASSO

---

**Input:** Response  $\mathbf{y}$ , Feature data  $X$ , Maximum degree  $d$ ,

Hyperparameter  $\lambda_1, \lambda_2$ .

**Output:** Feature index  $S$ , Term index  $S_t$ ,

Weights  $\hat{\beta}$  for  $d$ -degree polynomial basis.

1:  $\hat{\beta} = \arg \min_{\beta} (\frac{1}{2} \|\mathbf{y} - X\beta\|_2^2 + \lambda_1 \sum_{j=1}^{P+1} |\beta_j|)$

2:  $S = \{j : \hat{\beta}_j \neq 0\}$

3:  $X_{selected} = \Omega(X, S)$

4:  $X_{new} = \text{expan}(X_{selected}, d)$

5:  $w_j = (X_{new}^\top X_{new} + 0.001I)^{-1} X_{new}^\top \mathbf{y}$

6:  $\hat{\beta} = \arg \min_{\beta} (\frac{1}{2} \|\mathbf{y} - X_{new}\beta\|_2^2 + \lambda_2 \sum_j |\frac{\beta_j}{w_j}|)$

7:  $S_t = \{j : \hat{\beta}_j \neq 0\}$

---

Step3 の  $\Omega(X, S)$  は、元の行列から、インデックス  $S$  に含まれる特定の列だけ抽出した行列を得る関数である。たとえば、

$$S = \{0, 1, 2\}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N3} \end{bmatrix}, \quad (8)$$

のときには、

$$\Omega(X, S) = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} \end{bmatrix}, \quad (9)$$

となる。Step4 の  $\text{expan}$  関数は、説明変数に対して多項式展開を施す関数である。 $\text{expan}$  関数では、計画行列  $X$  と多項式展開する次数  $d$  を入力とし、元の計画行列  $X$  を  $d$  次分多項式展開した新たな行列を出力する。たとえば、 $P = 2$  の計画行列  $X$  を入力とし、 $d = 2$  と設定したときの  $\text{expan}$  関数の出力を以下に示す。

$$\text{expan}(X, 2) = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{11}x_{12} & x_{11}^2 & x_{12}^2 \\ 1 & x_{21} & x_{22} & x_{21}x_{22} & x_{21}^2 & x_{22}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N1}x_{N2} & x_{N1}^2 & x_{N2}^2 \end{bmatrix}. \quad (10)$$

ここで、 $\lambda_1$  は、ハイパーパラメータであり、学習データ内で交差検証を実施し、その中で二乗誤差が最小だった場

合のものを選定する。λ<sub>2</sub> もハイパーパラメータであり、学習データ内の交差検証で、赤池情報量基準 (AIC) [11] の下で最適となるものを選定する。上記のアルゴリズムにおいて、Step3 で取得した変数と残った変数の 2 乗や 3 乗の項、交互作用項の間には強い相関があるため、通常の LASSO を用いた場合には正しく変数の選択が行われないことがある [12]。そこで、Step6 では Adaptive-LASSO [13] を用いてパラメータの推定を行っている。

SPORE-LASSO を用いることによって、すべての変数に関する高次の項を考慮する必要がなくなり、目的変数の予測に対して有効な高次の項や交互作用項のみを対象とすることが可能になる。しかし、上記のアルゴリズムでは、あらかじめ展開する次数を定めておく必要がある。また、最大次数を大きく設定するにつれて、指数的にパラメータの数が増加してしまうという問題も抱えている。

### 3. 提案アルゴリズム

本稿では、多変量多項式回帰モデルのパラメータ  $\hat{\beta}$  の効率的な推定を目的とする。ただし、最大次数  $d$  は未知で  $\hat{\beta}$  は非ゼロのパラメータが少ないスパースなベクトルであるとする。

多項式回帰モデルのパラメータ推定を行う従来のアルゴリズムでは、最大次数を固定したうえで、推定が行われていた。一方で、本稿においては、最大次数が未知の場合におけるパラメータの推定アルゴリズムを提案する。

---

#### Algorithm 2 Proposed Method

---

**Input:** Response  $\mathbf{y}$ , Feature data  $\mathbf{X}$ , Hyperparameter  $\lambda_1, \lambda_2$ .

**Output:** Expanded feature data  $\mathbf{X}_{new}$ ,

Weights  $\tilde{\beta}$  for degree polynomial basis when converged.

- 1:  $\hat{\beta} = \arg \min_{\beta} (\frac{1}{2} \|\mathbf{y} - \mathbf{X}\beta\|_2^2 + \lambda_1 \sum_j |\beta_j|)$
  - 2:  $\xi = \|\mathbf{y} - \mathbf{X}\hat{\beta}\|_2^2 + \lambda_1 \sum_j |\hat{\beta}_j|$
  - 3:  $\mathbf{X}' = \mathbf{X}$
  - 4: Initialize  $\xi_{new}$  to a value smaller than  $\xi$
  - 5: **while**  $\xi_{new} < \xi$  **do**
  - 6:    $S = \{j : \hat{\beta}_j \neq 0\}$
  - 7:    $\mathbf{X}_{selected} = \Omega(\mathbf{X}', S)$
  - 8:    $\mathbf{X}_{new} = \text{multiply}(\mathbf{X}_{selected}, \mathbf{X})$
  - 9:    $\tilde{\beta} = \arg \min_{\beta} (\frac{1}{2} \|\mathbf{y} - \mathbf{X}_{new}\beta\|_2^2 + \lambda_2 \sum_j |\beta_j|)$
  - 10:    $\xi = \xi_{new}$
  - 11:    $\xi_{new} = \|\mathbf{y} - \mathbf{X}_{new}\tilde{\beta}\|_2^2 + \lambda_2 \sum_j |\tilde{\beta}_j|$
  - 12:    $\mathbf{X}' = \mathbf{X}_{new}$
  - 13:    $\hat{\beta} = \tilde{\beta}$
  - 14: **end while**
- 

はじめに、提案アルゴリズムで使用する関数の定義を行う。multiply( $\mathbf{X}', \mathbf{X}$ ) 関数は、 $\mathbf{X}$  の各列を  $\mathbf{X}'$  の各列に乗じて、新しい行列を取得する関数と定義する。たとえば、

$$\mathbf{X}' = \begin{bmatrix} x_{11}^2 & x_{11}x_{12} \\ \vdots & \vdots \\ x_{N1}^2 & x_{N1}x_{N2} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} \\ \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} \end{bmatrix}, \quad (11)$$

と与えられていたとすると、

$$\begin{aligned} & \text{multiply}(\mathbf{X}', \mathbf{X}) \\ &= \begin{bmatrix} x_{11}^2 & x_{11}x_{12} & x_{11}^3 & x_{11}x_{12}^2 & x_{11}^2x_{12} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{N1}^2 & x_{N1}x_{N2} & x_{N1}^3 & x_{N1}x_{N2}^2 & x_{N1}^2x_{N2} \end{bmatrix}, \end{aligned} \quad (12)$$

となる。上述の関数を用いて、最大次数が未知の多項式回帰モデルのパラメータを推定する提案アルゴリズムを以下に示す。

提案アルゴリズムの流れに関して説明する。まず、入力した計画行列  $\mathbf{X}$  に対して LASSO を適用する。次に、LASSO によって得られたパラメータの推定値  $\hat{\beta}$  を用いて、学習データ内における予測精度の評価値  $\xi$  を計算する。ここで、繰り返しの処理に入る前に、元の計画行列  $\mathbf{X}$  を  $\mathbf{X}'$  と置く。さらに、繰り返し計算が実行されるように評価値  $\xi_{new}$  には、 $\xi$  よりも小さい値の初期値を設定する。Step5 以降は繰り返しの処理を行うことを表している。提案アルゴリズムでは、繰り返しを実施する中で一時点前の予測精度の評価値  $\xi$  と現在の予測精度の評価値  $\xi_{new}$  を比較して収束判定を実施している。本稿では、現在の評価値  $\xi_{new}$  が悪化したときに収束したと判断する。

これから繰り返し処理の内容について具体的な例を用いながら説明を行う。まず、パラメータの推定値  $\hat{\beta}$  から非 0 と推定されたパラメータのインデックス  $S$  を抜き出す。インデックス  $S$  を用いて、計画行列  $\mathbf{X}'$  から一部の変数を取得する。上記の処理により、目的変数に有効な説明変数だけを持つ  $\mathbf{X}_{selected}$  という新たな行列を取得できる。

たとえば、元の計画行列  $\mathbf{X}$  から  $\mathbf{X}_{selected}$  が以下のよう

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N3} \end{bmatrix}, \quad \mathbf{X}_{selected} = \begin{bmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{N1} \end{bmatrix}. \quad (13)$$

次に、multiply 関数に  $\mathbf{X}_{selected}$  と元の計画行列  $\mathbf{X}$  を入力し、新たに  $\mathbf{X}_{new}$  という行列を得る。式 (13) を用いることで、以下のように  $\mathbf{X}_{new}$  を得ることができる。

$$\mathbf{X}_{new} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} & x_{11}^2 & x_{11}x_{12} & x_{11}x_{13} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N3} & x_{N1}^2 & x_{N1}x_{N2} & x_{N1}x_{N3} \end{bmatrix}. \quad (14)$$

multiply 関数によって、繰り返しを経るごとに、 $\mathbf{X}_{selected}$  には 1 つずつ次数が加わり、高次の多項式回帰モデルを構成できる。Step9 では、 $\mathbf{X}_{new}$  に対して、LASSO を適用

し、新たなパラメータの推定値  $\tilde{\beta}$  を得る。以降のステップで新しい予測精度の評価値  $\xi_{new}$  を算出するため、あらかじめ  $\xi$  に既存の  $\xi_{new}$  の値を代入している。Step11 では、 $X_{new}$  を用いて、Step2 同様に予測精度の評価値を計算し、 $\xi_{new}$  を得る。Step12–13 では、次の繰り返しで用いるために  $X_{new}$  を  $X'$  に置き換え、 $\tilde{\beta}$  を  $\hat{\beta}$  に置き換えている。

反復ステップを通じて、 $X_{new}$  がどのように変化していくか上記の例を用いて説明する。次の反復で、 $X_{selected}$  が以下のように得られたとする。

$$X_{selected} = \begin{bmatrix} 1 & x_{11}^2 \\ \vdots & \vdots \\ 1 & x_{N1}^2 \end{bmatrix}. \quad (15)$$

上記で得られた  $X_{selected}$  と元の計画行列  $X$  の各列を乗じる *multiply* 関数を適用することによって3次の項までを含む  $X_{new}$  を得る。

$$X_{new} = \begin{bmatrix} 1 & x_{11} & x_{12} & x_{13} & x_{11}^3 & x_{11}^2 x_{12} & x_{11}^2 x_{13} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{N1} & x_{N2} & x_{N3} & x_{N1}^3 & x_{N1}^2 x_{N2} & x_{N1}^2 x_{N3} \end{bmatrix}. \quad (16)$$

以上のように、収束するまで行列の更新を繰り返しつつ、評価値の比較を行っている。上記の処理により、 $X_{selected}$  の次数が1つずつ大きくなるため最大次数が未知の場合でも多項式回帰モデルに対する効率的なスパース推定を行うことが可能になる。

なお、予測を行う際には、収束の1つ前のステップで取得したパラメータの推定値  $\tilde{\beta}$  とその時の説明変数  $X_{new}$  を用いる。ただし、 $\lambda_1$  と  $\lambda_2$  はハイパーパラメータであり、学習データ内において、平均二乗誤差が最小になるように探索的に選択する。

#### 4. 人工データに対するシミュレーション実験

本章では、提案アルゴリズムの性能を評価するために、人工的に生成されたデータを用いてシミュレーション実験を行う。シミュレーション実験では、提案アルゴリズムを用いて、多項式回帰モデルに真のパラメータを設定して計算される目的変数をどれだけ正しく予測できるかについて検証する。

##### 4.1 実験条件

これから、人工データの生成法について詳細を示す。まず、説明変数は、標準正規分布に従って生成する。次に、上記で取得した説明変数に対して、*expan* 関数を適用し、多項式展開を施した説明変数を取得する。その後、新たに取得した説明変数にパラメータの真値を設定する。真のパラメータと多項式展開した説明変数を乗じて計算した値に、標準正規分布から生成した誤差項を加えたものを目的変数として扱う。ここで、真のパラメータは、多項式展開

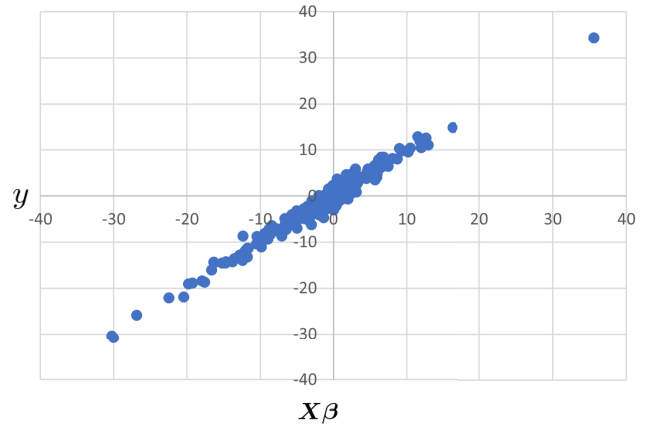


図 1 目的変数から誤差項の影響を除いた値と目的変数の値の散布図  
Fig. 1 Scatter plot of the target variable versus its error-excluded values.

した説明変数の列の長さと同じ数を持つ。なお、真のパラメータのうち20%だけに標準正規分布から生成した値を設定し、残りは0と設定している。また、今回の実験では、説明変数の数  $P = 4$ 、最大次数  $d = 4$  と設定して実験を行う。実際に実験で使用したデータの例を図1に示す。図1は、縦軸に目的変数、横軸に目的変数から誤差項を除いた値を描画している散布図である。図1より、両者の間の相関係数は0.9812であり、誤差項の影響は大きすぎるものではないことが分かる。

本研究の人工データによるシミュレーション実験と実データ実験には以下のようなPCを使用して行った。クアッドコアのIntel Core i7プロセッサで、16GBメモリを搭載し、オペレーティングシステムにはmacOS VenturaのMacBook Proを使用した。

次に実験の手順について示す。最初に、説明変数と真のパラメータを生成し、生成した説明変数と目的変数の組みを実験用データとして扱う。実験用データに対して、5分割交差検証を行い、データセット全体を訓練用データとテスト用データに分割し、訓練用データを用いてモデルを学習し、テスト用データに対して目的変数の予測精度とパラメータの推定精度を計算した。上記の手順を1,000回繰り返し、得られた結果を平均化して、提案アルゴリズムの性能を評価した。ただし、繰り返しごとに説明変数、真のパラメータ、誤差項は新たに生成している。サンプルサイズは50, 100, 150, 200, 250, 350, 450, 500の8パターンを設定した。

比較手法として、Random Forest [14]、SPORE-LASSO ( $d = 2, 3, 4, 5$ ) を使用した。なお、本研究では、提案アルゴリズムとSPORE-LASSOのスパース推定には、pythonの機械学習ライブラリscikit-learn [15]のバージョン1.2.1のLassoCV関数を用いている。スパース性を制御する引数alphasは、0.1から50まで設定しており、0.1刻みで探索を行った。

評価指標には、予測精度の観点で平均二乗誤差 (MSE) を用いる。

$$MSE = \frac{1}{M} \sum_{k=1}^M (t_k - \hat{t}_k)^2. \quad (17)$$

ここで、 $t_k$  は  $k$  番目のデータの実測値、 $\hat{t}_k$  は  $k$  番目のデータの予測値、 $M$  はデータ数を表す。パラメータの推定精度は、非 0 の真のパラメータのうち、推定したパラメータが含まれる割合 (Recall) と推定した非 0 のパラメータの中に、非 0 の真のパラメータが含まれる割合 (Precision) で評価する。真の非 0 パラメータ集合を  $S_0$ 、推定した非 0 パラメータ集合を  $\hat{S}_0$  とすると、

$$Recall = \frac{|S_0 \cap \hat{S}_0|}{|\hat{S}_0|}, \quad (18)$$

$$Precision = \frac{|S_0 \cap \hat{S}_0|}{|S_0|}, \quad (19)$$

のように表現できる。

### 4.2 結果と考察

まず、予測精度についての分析を行う。図 2 より、予測精度に関しては、サンプルサイズが大きくなるにつれて、提案アルゴリズムがあらかじめ定めた最大次数  $d = 4$  の SPORE-LASSO と同程度の予測精度を示すことが分かった。上記の結果から、最大次数を定めなくとも提案アルゴリズムを用いることによって、指定された次数のパラメータ推定が可能であると考えられる。しかし、提案アルゴリズム、SPORE-LASSO とともに、元のデータセット  $X$  に対するスパース推定の際に、定数項しかパラメータの推定が行われないケースが見受けられた。その結果、予測精度が著しく悪化しまう。上述の事象は、低次の項に真のパラメータが割り当てられず、高次の項だけに非 0 の真のパラメータが割り当てられている場合や偶数の次数にばかり真のパラメータが割り当てられた場合等のパターンで生じると考えられる。たとえば、図 3 のように偶数の次数で構成された偶関数は左右対称であるため傾きが捉えにくく、その結果として、 $x_1, x_2$  平面に平行になるよう定数項だけを推定してしまい、このような事象が発生してしまっている。上記のような定数項しか推定されない場合の実験結果を除いて、MSE を再計算した予測精度の比較結果が図 4 である。図 4 より、提案アルゴリズムは Random Forest に比肩するほどの予測精度を示していることが分かる。

次に、パラメータの推定精度について分析を行う。図 5、図 6 より、最大次数  $d = 4$  より小さい SPORE-LASSO の Recall は真のパラメータを含めることができない分、精度が良くなりやすいことが分かった。一方で、SPORE-LASSO の  $d = 5$  のような真のモデルを包含しているモデルは最大次数  $d = 4$  と同程度の精度が確認できた。他方、提案アルゴリズムもサンプルサイズが大きくなるにつれ、真の次数

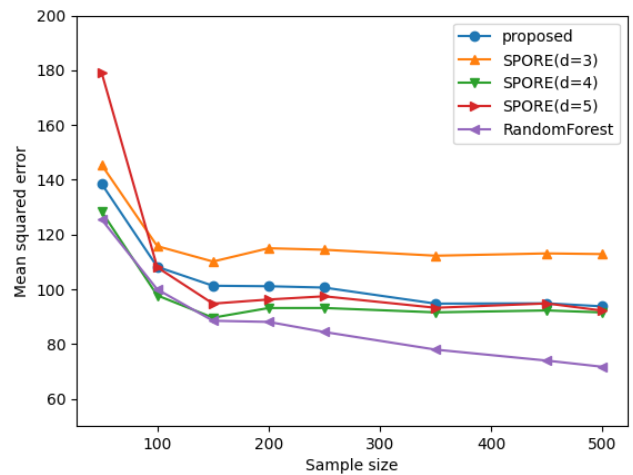


図 2 サンプルサイズと平均二乗誤差 (MSE) の推移

Fig. 2 Trends in sample size and Mean Squared Error (MSE).

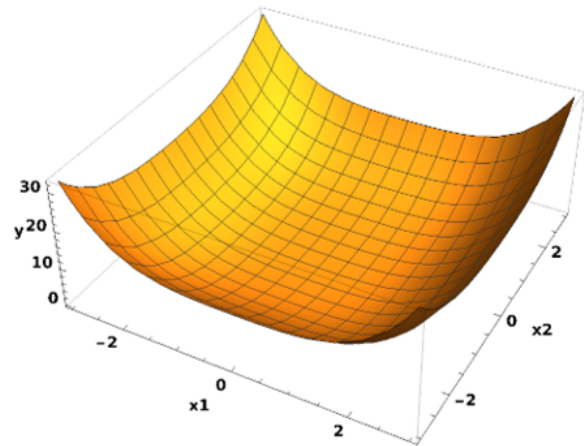


図 3 2 次の項と 4 次の項を含む多項式関数の 3 次元プロット

$x_1, x_2$ : 多項式関数の変数,  $y$ :  $x_1, x_2$  に対する関数の出力値

Fig. 3 3D plot of a polynomial function containing quadratic and quartic terms.  $x_1, x_2$ : Variables of the polynomial function,  $y$ : Output value of the function for the given  $x_1$  and  $x_2$ .

を設定した SPORE-LASSO と同程度の結果を得ることができている。図 7 から、提案アルゴリズムの方が  $d = 5$  の SPORE-LASSO よりも Precision の値が良くなっている。これらのことから、最大次数が未知であっても、提案アルゴリズムを用いることで、真のパラメータが割り当てられている項をある程度正しく推定できることが示された。SPORE-LASSO を用いる際には、大きく最大次数を設定することで良好な予測精度を得ることが可能と考えられるが、不要な項の推定を行ってしまう。その結果、余計な項が含まれている分、予測精度の劣化や構築したモデルの解釈性に影響を及ぼしてしまうことが考えられる。一方で、提案アルゴリズムは余計な項を含むことなく、最大次数が未知でありながら正しく真のパラメータを推定できていることから、解釈性の観点からも有用なアルゴリズムといえる。

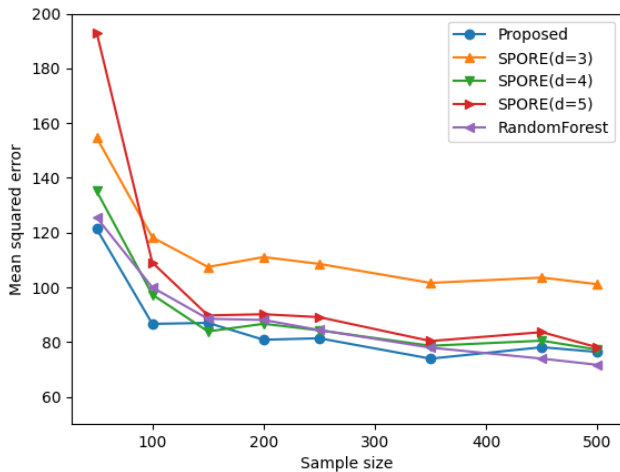


図 4 サンプルサイズと平均二乗誤差の推移  
(定数項のみ推定された場合を除去)

Fig. 4 Trends in Sample Size and Mean Squared Error (MSE)  
(Excluding where only the intercept was estimated).

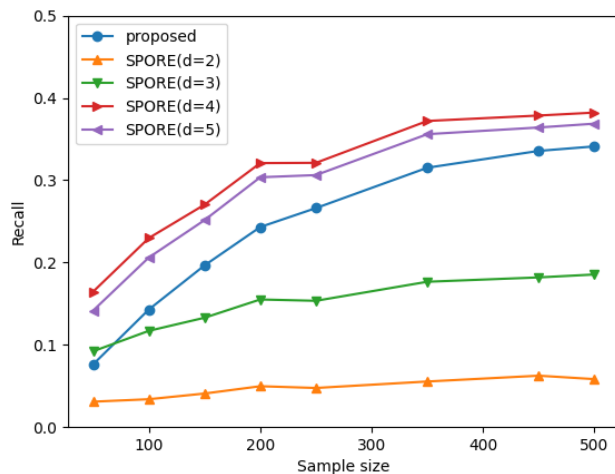


図 5 サンプルサイズと Recall の推移  
Fig. 5 Trends in sample size and Recall.

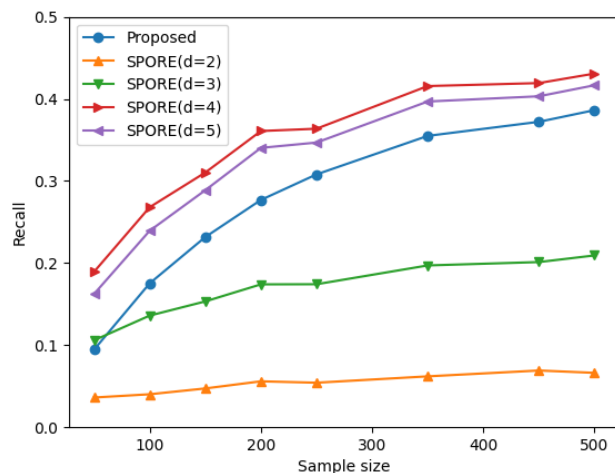


図 6 サンプルサイズと Recall の推移  
(定数項のみ推定された場合を除去)

Fig. 6 Trends in sample size and Recall  
(Excluding where only the intercept was estimated).

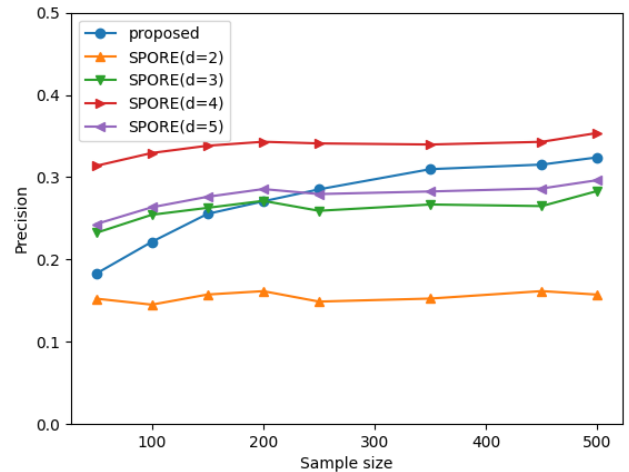


図 7 サンプルサイズと Precision の推移

Fig. 7 Trends in sample size and Precision.

表 1 提案アルゴリズムと SPORE-LASSO の 1 回あたりの平均実行時間の比較

Table 1 Comparison of average execution Time:  
Proposed Algorithm vs. SPORE-LASSO.

アルゴリズム	平均実行時間 (秒)	±標準偏差
Ours	0.6306	0.3681
SPORE( $d = 2$ )	0.2329	0.0546
SPORE( $d = 3$ )	0.2347	0.0556
SPORE( $d = 4$ )	0.2697	0.0831
SPORE( $d = 5$ )	0.3005	0.1071

他方, Recall と Precision の値の低さに関しては, LASSO を用いていることに起因していると考えている. 提案アルゴリズムでは, LASSO を用いて説明変数を選択できる点に着目し, 多変量多項式回帰モデルの予測フレームワークに活用している. しかし, LASSO は BIC [16] のように変数選択に関する理論的な保証がないため, 提案アルゴリズムにおいても選択された変数に対して一致性等の保証がない. その結果, 今回の実験で得られたような精度になったと考えられる. 今回は予測精度の向上のみに着目しているが, 変数選択に対する正当性や精度を担保するのであれば別のアルゴリズムを検討する必要があり, 今後の課題としたい.

さらに, 表 1 より実行時間の観点についても分析する. 提案アルゴリズムの計算時間と SPORE-LASSO の最大次数 2 から 4 までの計算時間の和を差し引くと, 提案アルゴリズムの方が 1 回あたり 0.1067 秒速いことが分かった. 故に, 従来手法で 1 つずつ最大次数の探索を行う場合には, 提案アルゴリズムを用いた方が効率的であると考えられる. しかし, 並列で複数の次数の SPORE-LASSO を動かした場合と比べると, SPORE-LASSO の方が効率的であると考えられる.

## 5. 実データ実験

### 5.1 実験条件

本稿では、提案アルゴリズムの予測性能を評価するために、実データによる実験も行った。実データによる実験では、UCI Machine Learning repository [17] のデータセットを用いて、提案アルゴリズムと比較手法の予測精度の比較検証を行う。

実験で用いるデータについては、表 2 のとおりである。7つの実データセットを使用して実験を行っている。次に、実験の手順について説明する。各データセットに対し、5分割交差検証を行った。交差検証により、データセット全体を訓練用データとテスト用データに分割し、訓練用データを用いてモデルを学習し、テスト用データに対しての予測精度を評価した。比較手法には、重回帰モデル、LASSO、Random Forest、LightGBM [24]、SPORE-LASSO ( $d = 2, 3, 4$ ) を使用した。評価指標には平均二乗誤差 (MSE) を用いる。

### 5.2 結果と考察

表 3 より実験の結果、Concrete Slump Test (SLUMP)・Concrete Slump Test (FLOW)・Computer Hardware のデータセットで提案アルゴリズムが他の手法よりも優れた精度を示した。上記のデータセットは、Random Forest や LightGBM といった決定木系の集団学習アルゴリズムの精度が悪く、連続的な関数を表現する提案アルゴリズムの有効性が確認できる。実際、Yeh の研究 [18] では、多項式回帰モデルを用いたモデリングが優れた精度を示していたことが報告されていた。また、Ein-Dor らの研究 [19] では、相対パフォーマンスを線形回帰モデルを用いて推定していた。これらのデータセットは、線形モデルで良い精度を示しており、多項式回帰モデルを用いる提案アルゴリズムでも相性が良いと考えられる。さらに、表 3 から上記データセットは、最大次数 4 と 6 の項が推定した多項式回帰モデルに含まれていたことが分かった。直観的には SPORE-LASSO の予測精度は、提案アルゴリズムで推定した最大次数と同じ時に精度が良くなると予想される。しかし、実際にはそれよりも大きい最大次数  $d = 5, 6$  を設定した時の方が良いことが確認された。SPORE-LASSO は最初のスパース推定時に一部の 변수を抜き出し、これらの変数だけで多項式回帰モデルを構築しているため、変数の組み合わせの選択肢が制限されてしまう。一方、提案アルゴリズムは、元の計画行列と乗じる処理を繰り返し行っていることから、変数の組合せが多様であるため予測精度が向上していると考えられる。

Yacht Hydrodynamics 以下のデータセットに関しては、LightGBM と RandomForest の性能が優れていた。その一方で、多項式回帰モデルを用いる提案アルゴリズムや SPORE-LASSO はいずれも予測精度が低くなっている。

表 2 検証で用いた実データセット一覧とその概要

Table 2 Validation datasets and descriptions

データセット名	概要	サンプルサイズ	変数の数
Concrete Slump Test [18]	コンクリートの取り扱い易さの程度を各種成分から予測	103	7
Computer Hardware [19]	CPU の相対パフォーマンスをメモリの量やサイクルタイム等から予測	209	7
Yacht Hydrodynamics [20]	ヨットへの抵抗を船体のデータ等から予測。実験を通じて取得された。	308	6
AUTO MPG [21]	車の燃料消費量を馬力や重さ等から予測	398	7
winequality-red [22]	赤ワインの官能評価値を pH やアルコール度数等から予測	1,599	11
winequality-white [22]	同上	4,898	11
Real estate valuation [23]	1 平米あたりの住宅価格を築年数や最寄り駅までの距離等から予測	414	6

Kummer ら [25] は、Yacht Hydrodynamics データを用い、AdaBoost で優れた精度を示したことを報告している。また、Gooljar らの研究 [26] では、winequality データセットや AUTO MPG においても、本研究同様に決定木系の機械学習アルゴリズムが優れた精度を示したことが報告されている。さらに、Yeh ら [23] は、ニューラルネットワーク [27] を用いた分析により、重回帰分析モデルよりもニューラルネットワークの方が予測精度が優れていることを示した。従来の研究でも触れられているように、上記のデータセットは、決定木系の機械学習モデルの予測精度が高くなることが知られている。一方、提案アルゴリズムや SPORE-LASSO といった多項式回帰モデルを用いるモデルとは相性が悪いことが考えられる。

ここで、提案アルゴリズムと比較手法間の実行時間の比較結果を表 4 から分析する。最大次数が 4 程度であれば、SPORE-LASSO の方が速いが、それ以上の次数を推定する場合には、提案アルゴリズムを用いた方が効率的であると言える。大きな次数を設定した場合には、SPORE-LASSO では変数の組み合わせの数が膨大になってしまい、計算量が増大する。それに対し、提案アルゴリズムでは、元の計画行列を乗じる処理を繰り返し行うことにより、効率的な推定が可能になったと考えられる。しかしながら、Yacht Hydrodynamics データや AUTO MPG データのように非常に高次元な項を推定する場合には、多大な時間を要してしまっており、さらなるアルゴリズムの効率化が今後の課題としてあげられる。



表 3 複数の実データセットに対する平均二乗誤差 (MSE)

Table 3 Mean Squared Error (MSE) across multiple real-world datasets

データセット名	Ours (Max Degree)	SPORE ( $d = 2$ )	SPORE ( $d = 3$ )	SPORE ( $d = 4$ )	SPORE ( $d = 5$ )	SPORE ( $d = 6$ )	Random Forest	LightGBM
Concrete Slump Test (SLUMP)	<b>47.876</b> (4)	53.928	54.699	55.585	51.022	52.281	50.207	66.756
Concrete Slump Test (FLOW)	<b>138.379</b> (6)	169.244	163.347	161.178	158.583	160.588	175.699	217.908
Concrete Slump Test (Strength)	6.075 (8)	2.480	1.514	1.153	0.888	<b>0.757</b>	13.482	14.462
Computer Hardware	<b>115.847</b> (4)	276.973	251.945	502.894	940.638	1720.744	1934.341	10114.852
Yacht Hydrodynamics	2.203 (22)	18.784	8.797	3.593	2.669	2.260	<b>1.033</b>	22.916
AUTO MPG	8.252 (20)	8.284	8.458	8.368	8.359	8.399	<b>7.346</b>	8.799
winequality-red	0.449 (5)	0.414	0.420	0.421	0.422	0.424	<b>0.335</b>	0.393
winequality-white	0.605 (12)	0.581	0.548	0.962	0.608	0.612	<b>0.375</b>	0.448
Real estate valuation	63.163 (4)	62.426	61.980	62.396	66.290	66.227	57.139	<b>56.199</b>

表 4 複数の実データセットに対する実行時間 (秒)

Table 4 Execution time (in seconds) across multiple real-world datasets.

データセット名	Ours	SPORE ( $d = 2$ )	SPORE ( $d = 3$ )	SPORE ( $d = 4$ )	SPORE ( $d = 5$ )	SPORE ( $d = 6$ )	Random Forest	LightGBM
Concrete Slump Test (SLUMP)	0.703	0.321	0.352	0.467	0.753	1.096	1.213	0.103
Concrete Slump Test (FLOW)	0.899	0.355	0.305	0.349	0.417	0.888	1.178	0.147
Concrete Slump Test (Strength)	0.961	0.328	0.322	0.864	1.600	3.092	1.165	0.367
Computer Hardware	3.535	0.398	0.495	1.868	3.720	7.286	1.324	0.799
Yacht Hydrodynamics	4.861	0.375	0.388	0.460	0.733	0.955	1.408	0.176
AUTO MPG	4.728	0.410	0.356	0.534	0.760	1.664	1.656	0.385
winequality-red	1.660	0.447	0.437	0.466	0.497	0.552	5.503	0.789
winequality-white	3.632	0.506	0.490	0.559	0.572	0.572	16.870	2.124
Real estate valuation	2.109	0.463	0.540	0.942	4.105	11.808	2.152	0.335

## 6. まとめと今後の課題

本稿では、最大次数が未知の場合でも多項式回帰モデルに対するスパース推定を可能にするアルゴリズムを提案した。従来の SPORE-LASSO は、最大次数が定まっている場合に効率的に多項式回帰モデルを可能にするアルゴリズムであった。しかし、このアルゴリズムは、最大次数があらかじめ分かっている必要があり、最大次数が定まらない場合には試行錯誤を要した。そのため、次数を大きく設定した際には計算量が大きくなってしまふ等の問題を抱えていた。

提案アルゴリズムは、学習データ内の精度が悪化するまで1つずつ次数を増やしながら行列を更新し、収束直前の多項式回帰モデルのパラメータと説明変数を出力する。上記の処理により、最大次数が未知の場合でも多項式回帰モデルのスパース推定を可能にした。シミュレーション実験の結果、提案アルゴリズムは、サンプルサイズが大きくなると、真の最大次数と同じ次数を設定した SPORE-LASSO と同等の予測精度を示すことが分かった。さらに、定数項のみしか推定されない場合の予測結果を除いて計算した予測精度では、Random Forest と同等の予測精度を示した。これらのことから、提案アルゴリズムは真の最大次数

に関する情報を必要とせずに、効率的に高次の項を推定できることが明らかとなった。また、実データ実験の結果、いくつかのデータセットにおいて提案アルゴリズムが優れた予測精度を示しており、実務でも使用できる可能性を示唆した。

提案アルゴリズムでは、アルゴリズムの初期段階での LASSO で定数項しか推定されない場合に精度が著しく悪化してしまうという問題を抱えている。こうした問題に対して、たとえば、 $d = 2$  の多項式展開を行い、強制的にアルゴリズムを進めるといった対処も考え得る。このような方法の有効性の検証は、今後の課題である。また、LASSO のみを用いているため変数の相関が強い場合に正しくパラメータを推定できていないとは限らず、SPORE-LASSO と同様に変数の相関を考慮できるよう Adaptive-LASSO の適用も考えられる。また、より複雑な関数系を示すために提案アルゴリズムによって、得られた関数に対して、何らかの非線形変換を施し、精度の向上を図る等の改良も考えられる。

謝辞 本研究の一部は、日本学術振興会科学研究費基盤研究 (C) 一般 (No.21K11796) により行われた。

参考文献

[1] Bishop, C.M.: パターン認識と機械学習 (上) ベイズ理論による統計的予測, Springer Japan (2008).

[2] Wang, G., Cao, J., Wang, H. and Guo, M.: Polynomial regression for data gathering in environmental monitoring applications, *Proc. IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pp.1307–1311 (2007).

[3] Ostertagova, E., Frankovský, P. and Ostertag, O.: Application of polynomial regression models for prediction of stress state in structural elements, *GJPAM*, Vol.12, No.4, pp.3187–3199 (2016).

[4] 増井紀貞: 機械学習の解釈性, *人工知能*, Vol.33, No.5, pp.641–644 (2018).

[5] Arrieta, A.B., Daz-Rodriguez, N., Del Ser, J., et al.: Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible, *Information Fusion*, Vol.58, pp.82–115 (2020).

[6] Shimizu, R., Megumi, M. and Masayuki, G.: An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information, *Knowledge-Based Systems*, Vol.239, p.107970 (2022).

[7] Rendle, S.: Factorization machines, *Proc. 2010 IEEE International Conference on Data Mining*, pp.995–1000 (2010).

[8] Huang, L., Jia, J., Yu, B., et al.: Predicting execution time of computer programs using sparse polynomial regression, *NeurIPS*, Vol.23, pp.883–891 (2010).

[9] Tibshirani, R.: Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B*, Vol.58, No.1, pp.267–288 (1996).

[10] Bertsimas, D. and Van Parys B.: Sparse hierarchical regression with polynomials, *Machine Learning*, Vol.109, No.5, pp.973–997 (2020).

[11] Sakamoto, Y., Ishiguro, M. and Kitagawa, G.: Akaike information criterion statistics, *Dordrecht, The Netherlands: D. Reidel*, Vol.81, No.10.5555, p.26853 (1986).

[12] Zhao, P. and Yu, B.: On model selection consistency of Lasso, *JMLR*, Vol.7, pp.2541–2563 (2006).

[13] Zou, H.: The adaptive lasso and its oracle properties, *Journal of the American Statistical Association*, Vol.101, No.476, pp.1418–1429 (2006).

[14] Breiman, L.: Random forests, *Machine Learning*, Vol.45, No.1, pp.5–32 (2001).

[15] Pedregosa, F., Varoquaux, G., Gramfort, A., et al.: Scikit-learn: Machine learning in Python, *JMLR*, Vol.12, pp.2825–2830 (2011).

[16] Neath, A.A. and Cavanaugh, J.E.: The Bayesian information criterion: Background, derivation, and applications, *WIREs Comp Stats*, Vol.4, No.2, pp.199–203 (2012).

[17] Asuncion, A. and Newman, D.: UCI machine learning repository, available from <https://archive.ics.uci.edu/> (accessed 2023-07-20).

[18] Yeh, I-C.: Simulation of concrete slump using neural networks, *Proc. Institution of Civil Engineers-Construction Materials*, Vol.162, No.1, pp.11–18 (2008).

[19] Ein-Dor, P. and Feldmesser, J.: Attributes of the performance of central processing units: A relative performance prediction model, *Comm. ACM*, Vol.30, No.4, pp.308–317 (1987).

[20] Keuning, L., JA., Onnink, R., Versluis, A., et al.: The bare hull resistance of the Delft systematic yacht hull series, *Proc. 14th International Symposium on Yacht*

*Design and Yacht Construction*, No.1102-P, pp.37–50 (1997).

[21] Quinlan, R.: Auto MPG Data Set, available from <https://archive.ics.uci.edu/ml/datasets/auto+mpg> (accessed 2023-07-14).

[22] Cortez, P. and Cerdeira, A., Almeida, F., et al.: Modeling wine preferences by data mining from physicochemical properties, *DSSs*, Vol.47, No.4, pp.547–553 (2009).

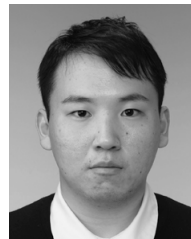
[23] Yeh, I-C. and Hsu, T-K.: Building real estate valuation models with comparative approach through case-based reasoning, *Applied Soft Computing*, Vol.65, pp.260–271 (2018).

[24] Ke, G., Meng, Q., Finley, T., et al.: Lightgbm: A highly efficient gradient boosting decision tree, *NeurIPS*, Vol.30, pp.3149–3157 (2017).

[25] Kummer, N. and Homayoun N.: Adaboost.MRT: Boosting regression for multivariate estimation, *JAIR*, Vol.3, No.4, pp.64–76 (2014).

[26] Gooljar, S., Manohar, K. and Hosein, P.: Performance Evaluation and Comparison of a New Regression Algorithm, arXiv preprint arXiv:2306.0910 (2023).

[27] Müller, B., Reinhardt, J. and Strickland, M.T.: Neural networks: An introduction, *Springer Science and Business Media* (1995).



井上 一磨

1995年生。2018年早稲田大学社会科学部卒業。2020年同大学大学院創造理工学研究科経営デザイン専攻修士課程修了。2020年株式会社エクサウィザーズ入社。機械学習モデルを用いた分析業務に従事。2023年早稲田大学大学院社会科学研究科政策科学論博士課程入学。



清水 良太郎

2017年早稲田大学創造理工学部経営システム工学専攻卒業。2019年同大学大学院修士課程修了。2023年同大学院博士課程修了。博士(工学)。2019～2020年株式会社ディー・エヌ・エーソフトウェアエンジニア。現在、ZOZO研究所リサーチサイエンティスト(2021年～)、カリフォルニア大学サンディエゴ校コンピュータサイエンス学部招聘研究員(2023年～)。



須子 統太 (正会員)

1978年生。2001年早稲田大学工学部経営システム工学科卒業。2003年同大学大学院修士課程修了。2006年同大学院博士課程単位取得退学。2009年博士(工学)取得。2005年早稲田大学工学部助手。2009年早稲田大学

メディアネットワークセンター助教。2014年早稲田大学社会科学総合学院専任講師。2016年同准教授。統計的学習理論、機械学習、機械学習の経営工学への応用に関する研究に従事。電子情報通信学会、日本経営工学会、各会員。



後藤 正幸 (正会員)

1969年生。1994年武蔵工業大学大学院修士課程修了。2000年早稲田大学大学院理工学研究科博士課程修了。博士(工学)。1997年同大学工学部助手。2000年東京大学大学院工学研究科助手。2002年武蔵工業大学環境情報学部助教授。2008年早稲田大学創造理工学部経営システム工学科准教授。2011年同大学教授。情報数理応用とデータサイエンス、ならびにビジネスアナリティクスの研究に従事。著書に、『入門パターン認識と機械学習』コロナ社(2014)、『ビジネス統計-統計基礎とエクセル分析』オデッセイコミュニケーションズ(2015)等。IEEE, INFORMS, 電子情報通信学会, 人工知能学会, 日本経営工学会, 経営情報学会等各会員。

2008年早稲田大学創造理工学部経営システム工学科准教授。2011年同大学教授。情報数理応用とデータサイエンス、ならびにビジネスアナリティクスの研究に従事。著書に、『入門パターン認識と機械学習』コロナ社(2014)、『ビジネス統計-統計基礎とエクセル分析』オデッセイコミュニケーションズ(2015)等。IEEE, INFORMS, 電子情報通信学会, 人工知能学会, 日本経営工学会, 経営情報学会等各会員。