パイナンバーを用いる外挿モデルの生成

枝光 敏章

ENJYN 株式会社 〒104-0061 東京都中央区銀座 7-13-6 サガミビル 2 階

E-mail: edamitsu.t@enjyn.co.jp

あらまし 一般に，機械学習モデルによる外挿は困難であり，これにより活用に制約が生じる．例えば新製品を開発する際，その新規性ゆえに，いずれかの説明変数（設計値や条件）が過去製品データの外挿領域に存在することが理由で，機械学習モデルが活用されないことがある．本論文は，パイナンバー（無次元量）を用いる外挿方法を提案する．本方法は，現象に対する人の知見や変数の単位を必要とせず，データからパイナンバーを自動生成する．パイナンバーが既知となれば，説明変数ベクトルが外挿領域に存在していても，パイナンバー変換・逆変換により，内挿領域にある説明変数ベクトルへ変換されれば，予測可能となる．円板の大たわみ現象に関するデータを用いた検証の結果，本方法により支配方程式に含まれるパイナンバーがデータから自動生成された．生成されたパイナンバーを用いたパイナンバー変換・逆変換により，外挿領域を含む領域に分布する説明変数ベクトルについて予測可能／不可能が判定された．予測可能と判定された説明変数ベクトルについては，内挿領域からの領域拡大の程度によらず，内挿と同等の精度で外挿が可能であることが示された．

キーワード 製品設計，外挿，パイナンバー，無次元量

1. はじめに

近年，効率化やイノベーション促進等を目的に，過去製品のデータから作成された機械学習モデルを対象にして多目的遺伝アルゴリズム等により最適な設計解を探索する手法[1]が活用されているが，新製品を開発する際，その新規性ゆえに，いずれかの説明変数（設計値や条件）が過去製品データの外挿領域に存在する場合がある．また，材料開発において，物質の過去にない組み合わせや条件に対する特性の予測が望まれている．しかし，従来のディープラーニング等の機械学習モデルによる外挿は困難である．また，機械学習モデルはブラックボックスとなるため現象の理解には貢献しない．

以上の様な課題を解決するために，データから支配方程式を特定する方法が報告されている．Schmidt と Lipson[2]は遺伝的プログラミングを用いる方法を提案した．Martius と Lampert[3]はニューラルネットワークを応用する方法 EQL を提案し，Sahoo ら[4]が改良を加えた．いずれの方法においても，候補となる演算子と関数が設定され，それらの組み合わせにより支配方程式が推定される．本研究では，支配方程式の形ではなく，支配方程式に含まれるパイナンバー（変数の乗べき積で表される無次元量）をデータから生成し，それらを用いる外挿モデル生成方法の構築を目指している[5][6]．本方法は，演算子や関数の候補を設定しないので，物理現象の種類や支配方程式の形に関する制約を受けない．

あらゆる現象に関する支配方程式はパイナンバーの関係式に変換可能であることがバッキンガムのパイ定理により示されている．通常，パイナンバーは人の作業により支配方程式や変数の単位（次元）から導かれる[7][8]が，本方法はデータのみからパイナンバーを自動生成する．パイナンバーは支配方程式に含まれる要素であるので，特定されたパイナンバーの形から現象理解が促進される可能性がある．

さらに，図 1 に示すように，外挿モデルと最適化アルゴリズムの連携により，既知のデータを用いて未知の領域に存在する最適解を探索する能力が得られる．これは，限られた経験から新しい問題に対応する，いわゆる応用力に相当する．

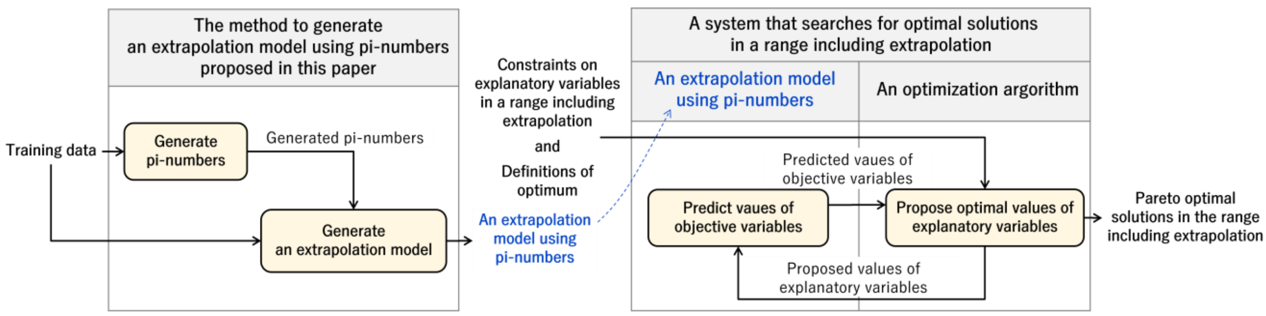


図 1 パイナンバーを用いる外挿モデルと最適化アルゴリズムの連携による応用力の獲得

以下に，パイナンバーを用いる外挿モデルの生成方法と外挿方法を提案し，円板の大たわみ現象に関するデータを用いた検証について述べる．

2. 物理量およびパイナンバーに関するデータ構造の定義

n個の物理量が関係する現象において，物理量をQで表すと，支配方程式は次式で表される．

$$f(Q_1, Q_2, \cdots, Q_n) = 0 \tag{1}$$

上式は，バッキンガムのパイ定理により，次式に変換される．

$$F(\Pi_1, \Pi_2, \cdots, \Pi_k) = 0 \tag{2}$$

物理量の持つ基本単位が 1 種類以上ある場合は，$k < n$となる． Πはパイナンバーを表し，以下の様に物理量の乗べき積で表される．

$$\Pi_i = {Q_1}^{p_{i,1}} {Q_2}^{p_{i,2}} \cdots {Q_n}^{p_{i,n}} \tag{3}$$

指数の$p_{i,1}, \cdots, p_{i,n}$は整数であり，次式でパイナンバー変換マトリックス$\boldsymbol{P}$およびパイナンバー変換ベクトル$\boldsymbol{p}$を定義する．

$$\boldsymbol{p_i} = (p_{i,1}, \cdots, p_{i,n})$$
$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{p_1} \\ \boldsymbol{p_2} \\ \vdots \\ \boldsymbol{p_k} \end{pmatrix} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,n} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,n} \\ \vdots & \vdots & & \vdots \\ p_{k,1} & p_{k,2} & \cdots & p_{k,n} \end{pmatrix} \tag{4}$$

特定の条件を表す物理量の値を要素に持つベクトルを物理量ベクトル$\boldsymbol{q}$，物理量説明変数ベクトルを$\boldsymbol{x}$，物理量目的変数をyと定義し，yは常に$\boldsymbol{q}$の第 1 要素に配置されるものとする．複数の$\boldsymbol{q}$を並べた 2 次元配列を物理量データセット$\boldsymbol{Q}$，物理量目的変数データセットを$\boldsymbol{Y}$，物理量説明変数データセットを$\boldsymbol{X}$と定義する．s個の条件がある場合，以下の様に表される．

$$\boldsymbol{q}_i = (q_{i,1}, q_{i,2}, \cdots, q_{i,n}) = (y_i, x_{i,1}, \cdots, x_{i,n-1}) = (y_i, \boldsymbol{x}_i)$$
$$\boldsymbol{Q} = \begin{pmatrix} \boldsymbol{q}_1 \\ \boldsymbol{q}_2 \\ \vdots \\ \boldsymbol{q}_s \end{pmatrix} = \begin{pmatrix} y_1 & \boldsymbol{x}_1 \\ y_2 & \boldsymbol{x}_2 \\ \vdots & \vdots \\ y_s & \boldsymbol{x}_s \end{pmatrix} = (\boldsymbol{Y}, \boldsymbol{X}) \tag{5}$$

パイナンバーベクトル$\boldsymbol{\pi}$，パイナンバー目的変数η，パイナンバー説明変数ベクトル$\boldsymbol{\xi}$，パイナンバーデータセット$\boldsymbol{\Pi}$，目的変数データセット$\boldsymbol{H}$およびパイナンバー説明変数セット$\boldsymbol{\Xi}$を以下の様に定義する．

$$\boldsymbol{\pi}_i = (\pi_{i,1}, \pi_{i,2}, \cdots, \pi_{i,n}) = (\eta, \xi_{i,1}, \cdots, \xi_{i,k-1}) = (\eta, \boldsymbol{\xi}_i)$$
$$\boldsymbol{\Pi} = \begin{pmatrix} \boldsymbol{\pi}_1 \\ \boldsymbol{\pi}_2 \\ \vdots \\ \boldsymbol{\pi}_s \end{pmatrix} = \begin{pmatrix} \eta_1 & \boldsymbol{\xi}_1 \\ \eta_2 & \boldsymbol{\xi}_2 \\ \vdots & \vdots \\ \eta_s & \boldsymbol{\xi}_s \end{pmatrix} = (\boldsymbol{H}, \boldsymbol{\Xi}) \tag{6}$$

複数の条件間で，$\boldsymbol{q}$が異なり，$\boldsymbol{\pi}$が同一である場合，それらの条件は物理的相似な関係となる．そのうちの 1 つの$\boldsymbol{q}$が既知となれば，他の物理的相似な現象の$\boldsymbol{x}$からyが求められる．

3. パイナンバー変換・逆変換

パイナンバー変換マトリックス$\boldsymbol{P}$を用いて$\boldsymbol{q}$を$\boldsymbol{\pi}$へ変換することを，パイナンバー変換と定義し，次式で表すことにする．

$$\boldsymbol{\pi} = f_\pi(\boldsymbol{q}, \boldsymbol{P}) \tag{7}$$

$\boldsymbol{P}$が既知であれば，$\boldsymbol{\pi}$は式(4)から決定される．$\boldsymbol{Q}$の中の全ての$\boldsymbol{q}$をパイナンバー変換することを次式で表すことにする．

$$\boldsymbol{\Pi} = f_\pi(\boldsymbol{Q}, \boldsymbol{P}) \tag{8}$$

パイナンバー変換マトリックス$\boldsymbol{P}$を用いて$\boldsymbol{\pi}$を$\boldsymbol{q}$へ変換することを,パイナンバー逆変換と呼ぶことにする．バッキンガムのパイ定理より，パイナンバーの個数は物理量の個数より少ないので，パイナンバー逆変換においては不定解となる.そこで,図 2 に示したように，解の領域を限定し，数値的に領域内の解を探索する．解が存在する場合，偶然的に収束した解が得られる．

パイナンバー変換マトリックス$\boldsymbol{P}$を用いて，物理量領域D内に$\boldsymbol{\pi}$をパイナンバー逆変換することを，次式で表すことにする．

$$\boldsymbol{q} = f_\pi^{-1}(\boldsymbol{\pi}, \boldsymbol{P}, D) \tag{9}$$

また，$\boldsymbol{\Pi}$の中の全ての$\boldsymbol{\pi}$をパイナンバー逆変換することを次式で表すことにする．

$$\boldsymbol{Q} = f_\pi^{-1}(\boldsymbol{\Pi}, \boldsymbol{P}, D) \tag{10}$$

次式のように，パイナンバー変換により$\boldsymbol{q}$を$\boldsymbol{\pi}$へ変換し，さらに物理量領域D内に$\boldsymbol{\pi}$をパイナンバー逆変換すると，$\boldsymbol{q}$と同一の$\boldsymbol{\pi}$を持つ（すなわち,物理的に相似な）$\boldsymbol{q}'$が得られる．

$$\boldsymbol{q}' = f_\pi^{-1}(f_\pi(\boldsymbol{q}, \boldsymbol{P}), \boldsymbol{P}, D) = f_\pi^{-1} \circ f_\pi(\boldsymbol{q}, \boldsymbol{P}, D) \tag{11}$$

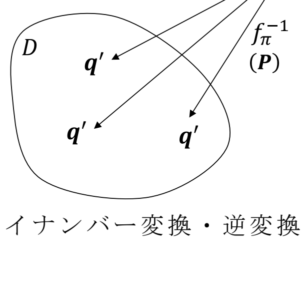


図 2　パイナンバー変換・逆変換の模式図

4. パイナンバー変換・逆変換を用いる外挿の方法

現象のパイナンバーが既知である場合，パイナンバー変換・逆変換により，データ分布領域$D(\boldsymbol{X})$の外にある物理量説明変数ベクトル$\boldsymbol{x}_{\mathrm{ext}}$と物理的に相似な$\boldsymbol{x}'_{\mathrm{ext}}$が$D(\boldsymbol{X})$内に見つけられれば，目的変数$y_{\mathrm{ext}}$の予測が可能となる．以下に処理の流れを述べる．

パイナンバー変換・逆変換を用いる外挿処理

(1) 学習用の物理量データセット$\boldsymbol{Q}$により回帰モデルf_{reg}を作成する．したがって，$D(\boldsymbol{X})$がf_{reg}の学習範囲となる．

(2) $\boldsymbol{P}$をq_1がΠ_1のみに含まれる形（$p_{2,1}$から$p_{k,1}$のすべてが 0）に変形する．

(3) パイナンバー変換・逆変換により，外挿となる説明変数$\boldsymbol{x}_{\mathrm{ext}}$と物理的に相似な$\boldsymbol{x}'_{\mathrm{ext}}$求める．

$$\boldsymbol{x}'_{\mathrm{ext}} = f_\pi^{-1} \circ f_\pi\left(\boldsymbol{x}_{\mathrm{ext}}, D(\boldsymbol{X})\right) \tag{12}$$

(4) 回帰モデルf_{reg}により，目的変数y'_{ext}を予測する．

$$y'_{\mathrm{ext}} = f_{\mathrm{reg}}(\boldsymbol{x}'_{\mathrm{ext}}) \tag{13}$$

(5) 目的変数が定まった$\boldsymbol{q}'_{\mathrm{ext}} = \left(y'_{\mathrm{ext}}, x'_{\mathrm{ext},1}, \cdots, x'_{\mathrm{ext},n-1}\right)$をパイナンバー変換することにより，$\boldsymbol{\pi}_{\mathrm{ext}}$を求める．この第 1 要素として$\eta_{\mathrm{ext}}$が得られる．

$$\boldsymbol{\pi}_{\mathrm{ext}} = f_\pi(\boldsymbol{q}'_{\mathrm{ext}}) \tag{14}$$

(6) π_1の定義式を変形した次式によりy_{ext}を求める．

$$y_{\mathrm{ext}} = \left(\frac{\eta_{\mathrm{ext}}}{x_{\mathrm{ext},1}{}^{p_{1,2}} x_{\mathrm{ext},2}{}^{p_{1,3}} \cdots x_{\mathrm{ext},n-1}{}^{p_{1,n}}}\right)^{\frac{1}{p_{1,1}}} \tag{15}$$

5. パイナンバー妥当性の評価方法

パイナンバー自動生成のためには，パイナンバーの妥当性を評価する指標が必要となる．ここでは，自己空間パイナンバー変換・逆変換法および回帰モデル法の 2 つを提案する．

5.1. 自己空間パイナンバー変換・逆変換法

物理量目的変数データセット$\boldsymbol{Y}$は既知であるが，パイナンバー変換・逆変換を用いる外挿の方法により，物理量説明変数データセット$\boldsymbol{X}$から予測値$\boldsymbol{Y}_{\mathrm{ext}}$を求めることもできる．$\boldsymbol{Y}_{\mathrm{ext}}$と$\boldsymbol{Y}$はパイナンバーが妥当であるほど近い値となる．それらが一致する程度を決定係数で評価し，${R^2}_\pi$と表す．ただし，予測値はf_{reg}の完成度の影響を受けるので，f_{reg}を作成した際の分割検証による検証スコア${R^2}_{\mathrm{model}}$で正規化した値をパイナンバー妥当性V_{trans}とみなす．

$$V_{\mathrm{trans}} = \frac{{R^2}_\pi}{{R^2}_{\mathrm{model}}} \tag{16}$$

以上のパイナンバー妥当性評価方法を自己空間パイナンバー変換・逆変換法と呼ぶことにする．

5.2. パイナンバー回帰モデル法

バッキンガムのパイ定理は，方程式で支配される現象には式(1)に示したようにパイナンバーの関係式Fが存在することを保証する．しかし，式中のパイナンバーが現象を表す正しいパイナンバーと異なっている場合は明確なFは存在しないので，機械学習が難しくなる．パイナンバーデータセット$\boldsymbol{\Pi}$を学習用および検証用データに分割し，前者により回帰モデルを作成し，後者により予測精度の検証をしたときの決定係数をパイナンバー妥当性V_{reg}とみなす．

$$V_{\mathrm{reg}} = R^2\left(\boldsymbol{H}_{\mathrm{test,pred}}, \boldsymbol{H}_{\mathrm{test}}\right) \tag{17}$$

ここで，$\boldsymbol{H}_{\mathrm{test,pred}}$および$\boldsymbol{H}_{\mathrm{test}}$は，それぞれ，検証用データにおける予測値および真値を表す．

以上のパイナンバー妥当性評価方法をパイナンバー回帰モデル法と呼ぶことにする．

6. パイナンバー自動生成方法

パイナンバー自動生成とは，データセットからパイナンバー妥当性が高いパイナンバー変換マトリックスを自動生成することである．そのために，パイナンバー変換マトリックスの提案とパイナンバー妥当性評価を繰り返して探索する方法が有効であると考えられる．パイナンバー妥当性評価方法については既に述べた．パイナンバー変換マトリックスの提案方法として，遺伝アルゴリズムが候補に挙がるが，組み合わせ数が多いため収束までの計算コストは高い．ここでは，パイナンバーの特性を利用して段階的にパイナンバー変換ベクトルを合成するパイナンバー提案方法について述べる．

6.1. 物理量をパイナンバーと同等に扱うパイナンバー変換ベクトルの合成

バッキンガムのパイ定理が示すパイナンバーの個数は次元解析から求められる個数であり，導出され得る個数の最大値である．次元解析から求められるパイナンバーを数学的パイナンバーと呼ぶことにする．一方，支配方程式あるいは物理法則の組み合わせから，より少ないパイナンバーが導出されることがある．それらは，基本単位の次元の整合性という数学的な必要条件を満たすだけでなく，物理的知見の活用により集約された結果であると解釈できる．このようなパイナンバーを物理的パイナンバーと呼ぶことにする．式(1)から導かれるパイナンバーの個数kは導出方法により

変わり，$n-1$個以下のいずれかの個数となる．

また，より少ない個数のパイナンバーが導かれる際，新たなパイナンバーは元のパイナンバーの乗べき積となる．すなわち，新たなパイナンバーは，物理量あるいはパイナンバーの乗べき積として合成される．

以上から，パイナンバー自動生成アルゴリズムの検討において，支配方程式に含まれる変量（式(1)における物理量，式(2)におけるパイナンバー，あるいは混在する物理量とパイナンバー）の区別は必要ではない．以下では，物理量もパイナンバーとして表現され，いずれの変量もパイナンバー変換ベクトルとして取り扱われる．

複数のパイナンバー変換ベクトルの線形結合は，それらが表すパイナンバーの乗べき積と等価である．すなわち，N個の変量を合成する場合，パイナンバー変換ベクトルの候補$\boldsymbol{p}_{\mathrm{new}}$は，それらを表すパイナンバー変換ベクトルの線形結合により作成される．

$$\boldsymbol{p}_{\mathrm{new}} = \sum_{i=1}^{N} C_i \boldsymbol{p}_i \tag{18}$$

ここで，C_iは整数であり，Π_iに対する指数を表す．

6.2. パイナンバー探索方法

式(18)の$\boldsymbol{p}_{\mathrm{new}}$が，パイナンバー変換マトリックス$\boldsymbol{P}$に加えられると新たなパイナンバー変換マトリックス$\boldsymbol{P}_{\mathrm{new}}$が提案される．提案された$\boldsymbol{P}_{\mathrm{new}}$のパイナンバー妥当性が高い場合，それを新たな$\boldsymbol{P}$として採用する．以上の繰り返しにより，段階的に正解のパイナンバーに近づいていく．その過程で，$\boldsymbol{p}_{\mathrm{new}}$の合成に使われた$\boldsymbol{p}_i$が削除されても高いパイナンバー妥当性が得られる場合は$\boldsymbol{p}_i$は削除される．

式(18)のC_iが取り得る値の範囲が広く設定されるほど，高次の物理量が提案されるが，組み合わせ数は増える．例えば，C_iが±1 の範囲に設定されると新たなパイナンバーの分子あるいは分母にΠ_iの 1 乗までの$\boldsymbol{p}_i$が含まれる候補に限定される．その際，もし正解の$\boldsymbol{p}_{\mathrm{new}}$に 2 乗以上の成分が含まれるなら，その途中の生成物である 1 乗までの組み合わせに限定された$\boldsymbol{p}_{\mathrm{new}}$が含まれる$\boldsymbol{P}_{\mathrm{new}}$に対して，提案前の$\boldsymbol{P}$よりも良いパイナンバー妥当性評価値が得られない可能性がある．この点については，円板のたわみ現象に関するデータによる検証結果において述べる．

初期のパイナンバー変換マトリックスとしては，物理量の個数と同じ大きさの単位行列が設定される．このときのパイナンバーは物理量そのものを表すので，物理量を元にしてパイナンバー自動生成が開始されることになる．

7. 検証用データセット

等分布荷重pを受ける円板のたわみ現象を対象として検証用データセットが作成された．図 3 は模式図を示す．現象に関係する物理量は，最大変位w_{max}，半径a，板厚h，等分布荷重pおよびヤング率Eである． w_{max}を目的変数，他を説明変数とする．大たわみ（非線形領域）を考慮する場合，次式が成り立つ[9].

$$\frac{w_{\mathrm{max}}}{h} + A\left(\frac{w_{\mathrm{max}}}{h}\right)^3 = B\frac{p}{E}\left(\frac{a}{h}\right)^4 \tag{19}$$

単純支持，半径方向の変位が自由である場合，$A = 0.262$，$B = 0.696$となる．表 1 は検証用データセットを示す． それぞれのデータセットの説明変数は，表に示した領域内にランダムな 125 条件の一様分布として作られた．データセット B から D の分布範囲は，データセット A の説明変数の最小値は維持され，最大値を大きくして設定された．外挿率R_{ext}はデータセット A に対する説明変数分布範囲の拡大率を表す．w_{max}の値は上式の数値解として求められた．パイナンバー自動生成はデータセット A，外挿検証はデータセット B，C，D を用いて実施された．以下の検証においては，それぞれの変数値をデータセット A の最大値で除することにより正規化された値$\tilde{w}_{max}, \tilde{p}, \tilde{a}\ \tilde{h}, \tilde{E}$が用いられた．

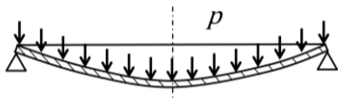


図 3　等分布荷重pを受ける円板のたわみ現象

表 1　円板のたわみ現象に関する検証用データセット

Data set	n	w_{max}, mm		p, Pa		a, mm		h, mm		E, GPa		R_{ext}
		min	max	min	max	min	max	min	max	min	max	
A	125	0.449	12.3	10	20	500	1500	1	3	50	100	1
B	125	0.065	29.6	10	30	500	2500	1	5	50	150	2
C	125	0.018	65.6	10	50	500	4500	1	9	50	250	4
D	125	0.013	146.6	10	90	500	8500	1	7	50	450	8

8. パイナンバー自動生成方法の検証

8.1. パイナンバー自動生成結果

パイナンバーの自動生成は，データセット A を用いて行われた．式(18)におけるC_iが取り得る値は-1，0，1 のいずれか，組み合わせられるパイナンバー変換ベクトルの数は$N = 2$に限定された．$\boldsymbol{p}_{\mathrm{new}}$の生成に$\boldsymbol{p}_1$が使われた場合は，$Q_1$が常に$\Pi_1$の分子のみに含まれるようにするために，以下のルールが加えられた．

$\boldsymbol{p}_{\mathrm{new}}$の生成に$\boldsymbol{p}_1$が使われた場合のルール

1.　C_1は 1 に固定される．

2. $\boldsymbol{p}_1$は常に$\boldsymbol{P}_{\text{new}}$から削除される．
3. 式(18)とは別に$p_{1,1}$に 1 が加えられた$\boldsymbol{P}_{\text{new}}$も候補に加えられる．

自己空間パイナンバー変換・逆変換法によるパイナンバー妥当性の合格基準は$V_{\text{thresh}} = 0.999$に設定された．

探索処理の検討により，パイナンバー変換・逆変換法は，正解のパイナンバーに対しては安定的に約 1.0 のV_{trans}を示すことがわかった．しかし，生成途中に部分的に正解に近づいた$\boldsymbol{P}_{\text{new}}$候補に対して，元のパイナンバーよりも高い$V_{\text{trans}}$を返すとは限らないことがわかった．すなわち，自己空間パイナンバー変換・逆変換法は，正解である際の判断は正確であるが，未完成な場合の差分を評価する精度は高くない．一方，パイナンバー回帰モデル法は，生成途中であっても説明変数の説明能力が増すと，その分だけパイナンバー空間における回帰モデルを作成しやすくなり，高いV_{reg}を示すことがわかった．以上を考慮して，以下の探索処理が構築された．

<u>パイナンバー探索処理</u>

(1) 上記条件の範囲の全ての組み合わせで$\boldsymbol{P}_{\text{new}}$候補を生成する．
(2) パイナンバーの数が減る$\boldsymbol{P}_{\text{new}}$候補と$p_{1,1}$に 1 が加えられた$\boldsymbol{P}_{\text{new}}$候補（上記$\boldsymbol{p}_{\text{new}}$の生成に$p_1$が使われた場合のルール 3）のパイナンバー妥当性$V_{\text{trans}}$を評価し，$V_{\text{trans}} \geq V_{\text{thresh}}$となる$\boldsymbol{P}_{\text{new}}$候補のうち，$V_{\text{trans}}$が最大の$\boldsymbol{P}_{\text{new}}$を採用する．$V_{\text{trans}} \geq V_{\text{thresh}}$となる$\boldsymbol{P}_{\text{new}}$がない場合は，パイナンバーの数が減らない$\boldsymbol{P}_{\text{new}}$のパイナンバー妥当性$V_{\text{reg}}$を評価し，$V_{\text{reg}}$が最大の$\boldsymbol{P}_{\text{new}}$を採用する．
(3) パイナンバーの個数が 2 となる，あるいは$V_{\text{trans}} \geq V_{\text{thresh}}$となる$\boldsymbol{P}_{\text{new}}$が現れなくなるまで以上を繰り返す．

パイナンバー逆変換における数値解の探索，内挿領域に対する回帰モデル作成およびパイナンバー回帰モデル法における回帰モデル作成には，それぞれ，Python のオープンソースパッケージである SciPy の L-BFGS-B[10]，scikit-learn の MLPRegressor [11] および RandomForestRegressor [12]が用いられた．

パイナンバー自動生成の結果，7 ステップ目で次のパイナンバー変換マトリックスが得られた．

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 4 & 4 & 1 \end{pmatrix} \tag{20}$$

図 4 は，各ステップで採用された$\boldsymbol{P}_{\text{new}}$のパイナンバー妥当性の履歴を示す．ステップ 0 は初期状態を表す．ステップ 1 以降，$V_{\text{trans}} \geq V_{\text{thresh}}$となる$\boldsymbol{P}_{\text{new}}$候補が存在しないとき，$V_{\text{reg}}$を徐々に向上させる$\boldsymbol{P}_{\text{new}}$が生成されている．

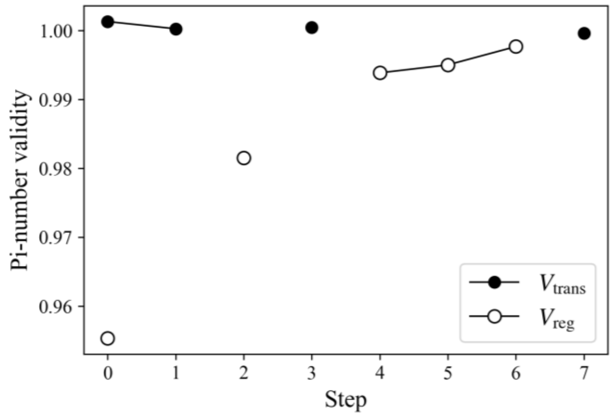


図 4　パイナンバー探索処理において採用されたパイナンバー変換マトリックスのパイナンバー妥当性の履歴

8.2. 自動生成されたパイナンバーと理論式から導かれるパイナンバーの比較

表 2 は，式(19)に含まれる物理量と基本単位の次元を示す．基本単位は N および m とした．

表 2　物理量と基本単位の次元

	$w_{\max}$	p	a	h	E
N	0	1	0	0	1
m	1	-2	1	1	-2

次元解析を実施すると，3 個の数学的パイナンバーが得られた．

$$\Pi_{\text{m}} = \{\Pi_{\text{m},1}, \Pi_{\text{m},2}, \Pi_{\text{m},3}\} = \left\{\frac{w_{\max}}{h},\ \frac{p}{E}, \frac{a}{h}\right\} \tag{21}$$

パイナンバーを要素に持つ集合をΠで表し，パイナンバーセットと呼ぶことにする．ここでは，数学的パイナンバーを表すために添え字 m を付した．次式のように，支配方程式である式(19)に，Π_{m}の要素が含まれていることがわかる．

$$\Pi_{\text{m},1} + A\Pi_{\text{m},1}{}^3 = B\Pi_{\text{m},2}\Pi_{\text{m},3}{}^4 \tag{22}$$

次元解析の際，基本単位の数は物理量の数よりも少ないので方程式は閉じない．そのため，パイナンバーの形は 1 つに定まらない．例えば，Π_{m}は，以下の形等に変形可能である．

$$\left\{\frac{w_{\max}}{h},\ \frac{p}{E}, \frac{a}{h}\right\} = \left\{\frac{w_{\max}}{a},\ \frac{p}{E}, \frac{a}{h}\right\} = \left\{\frac{w_{\max}}{a},\ \frac{E}{p}, \frac{w_{\max}}{h}\right\} \tag{23}$$

この変形は，パイナンバー変換マトリックスにおける行基本変形を施すことに相当する．この変形により，説明変数が 1 つのパイナンバーのみに含まれる形に必

ず変形できる．その操作は，パイナンバー変換マトリックスにおいて，物理量説明変数を左端の列に置いた上で階段行列を作る操作に相当する．以下では，パイナンバー変換マトリックス間の比較を容易にするために，パイナンバー変換マトリックスの表記を簡約化した形に統一する．

Π_mはパイナンバー変換マトリックスとしては次式で表現される．

$$\boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 \end{pmatrix} \tag{24}$$

上式が表す数学的パイナンバーは，図 4 の 3 ステップ目に生成されていた．

7 ステップ目で生成された式(20)が表すパイナンバーは，数学的パイナンバーよりも 1 個少なく，式(19)の右辺の物理量を一まとめにして認識することにより導かれる以下の物理的パイナンバーセットと同一であった．

$$\Pi_\mathrm{p} = \{\Pi_\mathrm{p,1}, \Pi_\mathrm{p,2}\} = \left\{\frac{w_\mathrm{max}}{h},\ \frac{pa^4}{Eh^4}\right\} \tag{25}$$

上式を用いて，式(19)は次式で表現される．

$$\Pi_{p,1} + A{\Pi_\mathrm{p,1}}^3 = B\Pi_{p,2} \tag{26}$$

従来は物理的パイナンバーの導出には物理的知見の活用が必要であったが，本検証においては，データから物理的パイナンバーが導出された．

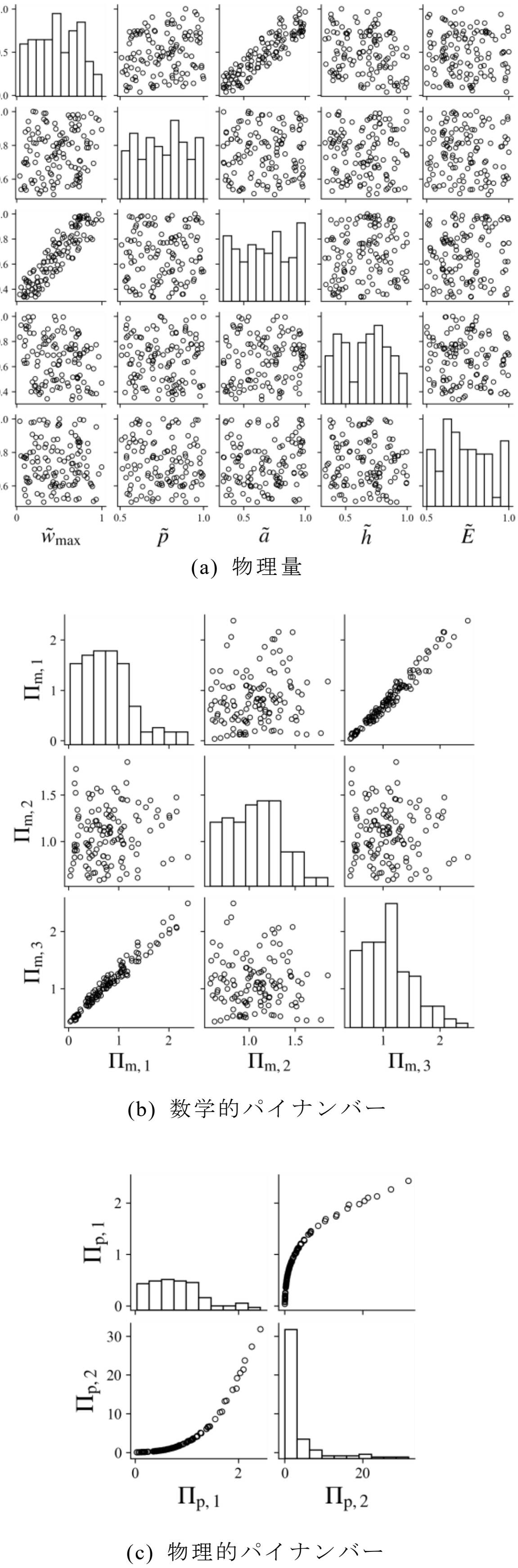
9. パイナンバーを用いる外挿方法の検証

9.1. パイナンバー変換による目的変数の説明変数への依存性の変化

図 5 は，データセット A の(a)物理量，(b)数学的パイナンバーおよび(c)物理的パイナンバーの散布図マトリックスを示す．視覚的に，物理量では$\tilde{a}$，数学的パイナンバーでは$\Pi_\mathrm{m,3}$，物理的パイナンバーでは$\Pi_\mathrm{p,2}$が，それぞれの説明変数の中で最も目的変数との関係が明確である．また，この順に関係が明確になっており，より少ない変数に合成されるほど，説明変数の説明能力が向上していることが示されている．特に，図(c)においては，式(26)のパイナンバーの関係式が明確に認識されているので，式(26)が未知であってもデータの分布から回帰式として式(26)が導かれる可能性がある．その際は，式(26)により，どのような物理量説明変数ベクトルに対しても外挿が可能となるので，機械学習モデルは不要となる．

9.2. パイナンバーを用いる外挿方法と通常の回帰モデルの比較

予測対象条件数のうち予測可能となる条件数の割



(a) 物理量

(b) 数学的パイナンバー

(c) 物理的パイナンバー

図 5　データセット A の散布図マトリックス

合を予測可能割合R_{pred}で表すことにする．図 6 は，外挿率R_{ext}を変化させたときの，パイナンバーを用いた外挿モデルおよび従来の物理量を用いた回帰モデルへの影響を示す．凡例中のnは変量の個数を表す．各点はデータセット A から D を用いて計算された． 図 6 (a)は R_{pred}の変化を表している．通常の回帰モデルのR_{pred}は説明変数が 4 つのときの理論式$(1/R_{\mathrm{ext}})^4$による計算値ある．パイナンバーによるR_{pred}は通常の回帰モデルよりも大きく，変数の個数が少ないほど大きい．図 6 (b)は予測値と真値との決定係数R^2の変化を表している．外挿率が増加するほど，通常の回帰モデルによる決定係数は低下していったが，パイナンバーを用いる外挿では，大きな変化はなかった．

以下に，一例として，データセット C（$R_{\mathrm{ext}}=4$）に関する結果について詳細を述べる．図 7 は，物理的パイナンバーを用いてデータセット A の内挿領域にパイナンバー変換・逆変換されたデータセット C の散布図マトリックスを示す．元のデータセット C は一様分布であるが,図中の点は偶然的に収束した解であるので，分布に密度の偏りがみられる．通常の回帰モデルが信頼できる内挿領域の割合はR_{pred} =0.00039（0.039 %）であるが，物理的相似な条件が見つかった条件は 107 個であったので，物理的パイナンバーを用いると$R_{\mathrm{pred}}=$ 107/125=0.856（85.6 %）が予測可能と判定された．

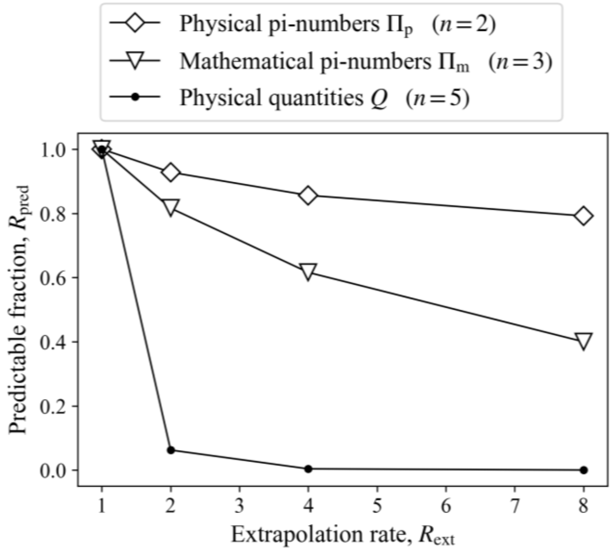
図 8 はデータセット C に対する予測値と式(19)の数値解の比較を示す．図 8 (a)は，全 125 条件に対する物理量回帰モデルによる予測結果である．決定係数は$R_{\mathrm{ext}}=1$のときの 1.0 から 0.864 に悪化している．図 8 (b)は，物理的パイナンバーによるパイナンバー変換・逆変換が可能であった 107 条件に対する外挿結果である．決定係数は，$R_{\mathrm{ext}}=1$のときの 0.999 から大幅な悪化はなく，0.992 であった．

10. まとめ

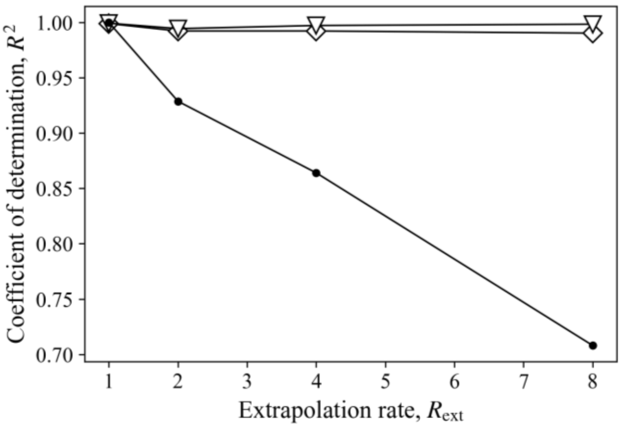
支配方程式の構成要素であるパイナンバー（無次元量）をデータから自動生成し，それらを機械学習と組み合わせることにより，外挿する方法が提案された．検証は円板の大たわみ現象に関するデータを用いて実施された．

従来，物理的パイナンバーの導出には人の物理的知見と作業が必要であったが，自己空間パイナンバー変換・逆変換法およびパイナンバー回帰モデル法を併用するパイナンバー自動生成方法により，データから物理的パイナンバーが生成された．メカニズムが不明な現象に関するデータに本方法が用いられると，自動生成されたパイナンバーにより現象の解明が促進される可能性がある．

パイナンバーを用いる外挿方法は，外挿領域にある



(a) 予測可能率



(b) 決定係数

図 6　外挿率を変化させたときの
パイナンバーを用いた外挿モデルおよび
従来の物理量を用いた回帰モデルへの影響

説明変数ベクトルについて予測可能であるか否かの判定が可能であることが示された．これにより，例えば製品設計において，外挿領域にある設計案に対して性能予測は不可能と判定された場合でも，予測可能となる設計値への変更検討が可能となる．外挿可能となる確率はパイナンバーの個数が少ないほど高くなり，予測可能な場合は外挿率によらず予測精度は内挿とほぼ同等であることが示された．

図 1 に示したように外挿モデルと探索アルゴリズムが連携すれば，予測可能であり且つ設計目標を満足する設計値が外挿領域に発見される可能性がある．

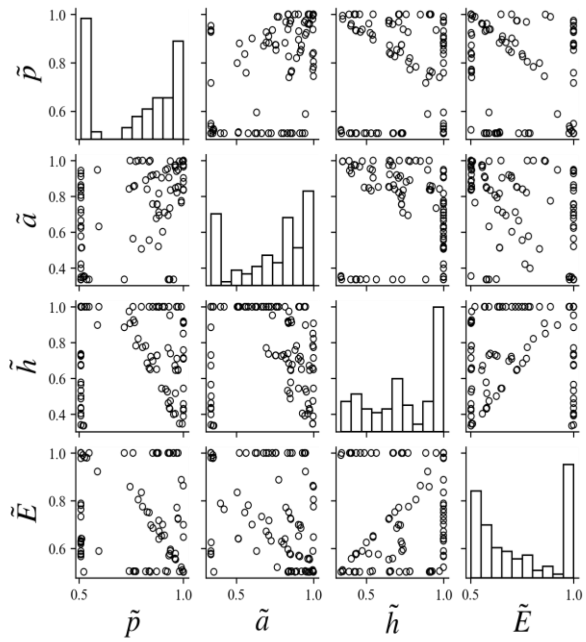
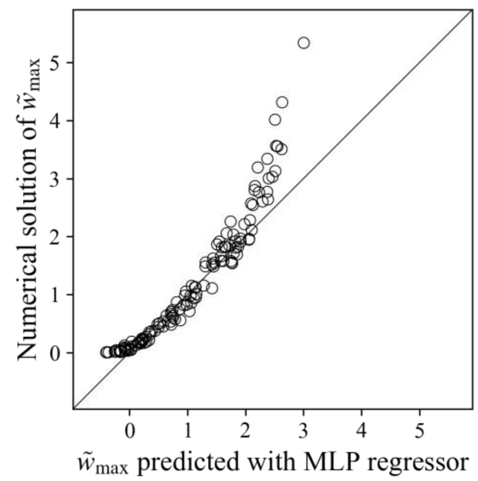
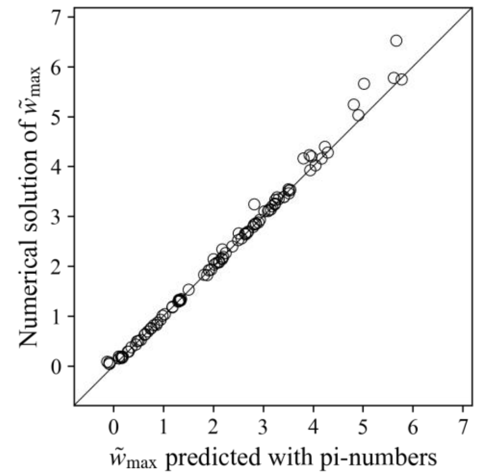


図 7　物理的パイナンバーによりデータセット A の内挿領域にパイナンバー変換・逆変換された データセット C の説明変数の散布図マトリックス （107 条件）



(a) 物理量の回帰モデルによる予測（125 条件）



(b) 物理的パイナンバーを用いる外挿 （パイナンバー変換・逆変換が 可能であった 107 条件）

図 8　データセット C に対する予測値と 数値解の比較

参　考　文　献

[1] 枝光敏章．機構の構想設計　シミュレーションモデル統合と最適化によるイノベーションの促進．機械設計 第 63 巻　第 10 号，日刊工業新聞社，pp. 42-47, 2019.

[2] Schmidt, M. and Lipson, H. Distilling free-form natural laws from experimental data. *Science* 324, pp. 81–85, 2009.

[3] Martius, G. and Lampert, C. H. Extrapolation and learning equations. *arXiv preprint arXiv:1610.02995*, 2016.

[4] Sahoo, S. S., Lampert, C. H., and Martius, G. Learning Equations for Extrapolation and Control. *Proceedings of the 35th International Conference on Machine Learning*, PMLR 80, 2018.

[5] ENJYN 株式会社．枝光敏章．データ解析方法，データ解析装置，及び，データ解析プログラム．特許第 6999207 号，2021.

[6] ENJYN 株式会社．枝光敏章．データ解析方法，データ解析装置，及び，データ解析プログラム．特許第 7039090 号，2022.

[7] 枝光敏章，柳瀬治，赤瀬賢一，西澤和毅．組織的 CAE 推進のためのシステム思考とパイナンバーの活用．日本機械学会 2020 年度年次大会講演論文集，J12104，2020.

[8] 江守一郎，斉藤孝三，関本孝三．模型実験の理論と応用　第三版．技報堂出版，2015.

[9] 日本機械学会編．機械工学便覧 DVD-ROM 版．材料力学　第 5 章　平板の曲げ，α 3-55, 丸善出版，2018.

[10] “scipy.optimize.minimize”, SciPy documentation, https://docs.scipy.org/doc/scipy/reference/optimize.minimize-lbfgsb.html, 12 Feb 2023.

[11] “sklearn.neural_network.MLPRegressor”, scikit-learn Machine Learning in Python, https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html, 12 Feb 2023.

[12] “sklearn.ensemble.RandomForestRegressor”, scikit-learn Machine Learning in Python, https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html, 12 Feb 2023.