

機械学習による空力解析手法の開発およびその検証

Development and verification of aerodynamic analysis method by machine learning

○ 豊田 耀山, 東北大院情報科学・流体研, 〒980-8577 仙台市青葉区片平 2-1-1,

E-mail: toyota@dragon.ifs.tohoku.ac.jp

服部 裕司, 東北大流体研, 〒980-8577 仙台市青葉区片平 2-1-1, E-mail: hattori@ifs.tohoku.ac.jp

Akitaka TOYOTA, GSIS & Institute of Fluid Science, Tohoku Univ., Aoba-ku Sendai 980-8577, JAPAN

Yuji HATTORI, Institute of Fluid Science, Tohoku Univ., Aoba-ku Sendai 980-8577, JAPAN

Prediction of the aerodynamic forces acting on an object in the external flow is performed by a neural network. The neural network is trained by NACA 4-digit wings, and verified by NACA 5-digit wings of which data are obtained by the 2-dimensional incompressible inviscid flow simulation. Efficient learning has been achieved by object shape representation using Fourier sine coefficients and reduction of the number of data sets by clustering. As a result, aerodynamic analysis method for airfoils in 2-dimensional incompressible inviscid flow has been established.

1. 緒言

空力解析は、航空機や橋梁などの設計において性能や強度を評価するために広く活用されている。

しかしながら空力解析は、その過程で流体の方程式を解くため計算量が多く、(1)大規模な計算設備が必要、(2)膨大な計算時間が必要、という問題を抱えている。そこで、精度を維持しつつ、計算量を削減する手法が多数提案されてきた⁽¹⁾。しかし、既往研究では理論式の変形や、数値計算の収束性改善に主眼を置いており、根本的な計算量の削減は達成できなかった。

そのため本研究では、機械学習の手法を応用し、流体の方程式を解かない空力解析手法を構築することに焦点を当てる。本稿では、このアイデアを実現するための第1段階として、2次元非圧縮性非粘性流を対象に、教師データの生成方法、物体形状の離散化と、教師データの効率的な削減方法に関して検討する。

2. 計算手法

2. 1. 機械学習の適用方法

未知関数 G について以下の方程式を仮定する。

$$\mathbf{y} = G(\mathbf{x}, \alpha, Ma, Re) \quad (1)$$

ここで、 \mathbf{y} は空力特性を並べた目的変数であり、 \mathbf{x} は、 \mathbf{s} 、 α 、 Ma 、 Re から構成される説明変数である。これらはそれぞれ、 \mathbf{s} :物体の形状、 α は迎角、 Ma はマッハ数、 Re はレイノルズ数である。このとき、 \mathbf{x} は定常状態の流れ場を決定するために必要十分なパラメータを含んでいることから、任意の入力 \mathbf{x} に対応する出力 \mathbf{y} が存在し、 $G(\mathbf{x})$ を計算することで空力解析が実行できる。

もし、 G をニューラルネットワーク (NN) によって近似できた場合、 G は有限回の行列の乗算に置換される。この計算量は一般的なFVMスキームの時間発展1ステップ分と同じオーダーであることから、圧倒的な高速化が期待できる。また、教師データとして用いる (\mathbf{x}, \mathbf{y}) の組は、CFD解析から作成できるほか、実験結果を利用することもできるため、拡張性に富む。

2. 2. 実装上の課題

実装上の主な課題は次の2つである。(a):式(1)中の物体形状 \mathbf{s} をどう表現すべきか、(b):教師データを生成するためのCFD解析の回数をどう減らすのかである。

課題 (a) は、物体形状を関数と考え、無限次元データを低次元データへ圧縮する問題と解釈できる。そのため、(a1) 点列の座標を等間隔に並べる手法、(a2) a1にて曲率が大きい部分に点を集中させる手法、(a3) Fourier 正弦係数を用いる手法、のそれぞれ3

つの物体形状表現を実装し、検証データに対する精度を決定係数: R^2 にて比較検討する。

ここで Fourier 正弦係数による物体形状表現の詳細を述べる。物体形状が、弧長パラメータ t 表現の閉曲線: $f(x(t), y(t)) = 0$ で与えられたとする。このとき、 $x(t), y(t) \in R, (t \in [0, 1])$ は常に一価関数であるから Fourier 級数展開で表現可能である。そこで、 x, y を $[-1, 1]$ で周期2の奇関数として展開し式(2)を得る。

$$f(x, y) \sim f\left(\sum_{k=1}^{\infty} b_{x_k} \cdot \sin(k\pi t), \sum_{k=1}^{\infty} b_{y_k} \cdot \sin(k\pi t)\right) \quad (2)$$

このとき、有限の d 項までの Fourier 正弦係数 b_{x_i}, b_{y_i} を昇順に並べた $2d$ 次元ベクトルを物体形状 \mathbf{s} とすれば、 \mathbf{s} から連続値を復号できるため、高い解像度を確保できる。このとき \mathbf{s} は式(3)の形となる。

$$\mathbf{s} = [b_{x_1}, b_{x_2}, \dots, b_{x_d}, b_{y_1}, b_{y_2}, \dots, b_{y_d}]^T \quad (3)$$

課題 (b) については、機械学習を内挿処理とみなすと、教師データは参照点であるから、類似性の高いデータを削除しても精度は維持できると考えられる。それを確認するため、今回は大量の NACA 翼型を用いて学習を行い、(b0) データ全体、(b1) 迎角のパターンを等間隔に減らしたもの、(b2) 翼型の種類を翼の番号に基づき等間隔に減らしたもの、(b3) クラスタリングにより類似した翼型を除去したもの、(b4) 類似した翼型を集めたもの、という5つのデータセットを作成し、検証データに対する精度を R^2 で比較検討する。

次にクラスタリングによって類似したデータを除去するアイデアの概要を示す。

多数の翼型を含むデータセットを N 個の d 次元ベクトル \mathbf{x}_i からなる集合 $X = \{\mathbf{x}_i \in R^d | 1 \leq i \leq N\}$ と記述するとき、各データ \mathbf{x}_i は、 R^d 空間中の点である。式(2)および(3)より、類似した翼型は、 R^d 空間中で距離が近い点になると期待される。そこで X が、近傍の点のみで構成された複数の部分集合 C_j の和集合であると仮定し、適当なクラスタリング手法で X を M 個の C_j に分割する。各 C_j に割り当てられた \mathbf{x}_i を m 個ずつ代表値として抽出することで、 N 個のデータから類似性の低い $m \cdot M$ 個のデータセットが作成できる。

この手法のうち変更可能な要素は、クラスタリング手法、データ前処理法、部分集合 C_j の代表値の決定方法の3つである。本研究では、クラスタリング手法として k -means⁺⁺(2)を用い、前処理は行わず、 C_j の代表値は C_j の重心に最も近いデータを採用した。

2. 3. 教師データと使用したNNの構成

用意した教師データ・テストデータの諸元を Table 1 に示す。

Table 1 Variables patterns & total number of dataset.

Stage	type	α	s	Total($\alpha \times s$)
Training	NACA4digit	40	5000	200000
Validation	NACA5digit	40	891	35640

今回は、非圧縮非粘性流を扱うため、 Re と Ma とを変数から除外したうえで、迎角は -40° から 38° まで 2° 刻みで40パターン用意した。物体形状は、学習用にNACA4桁翼の奇数番のみ5000種類、検証用にNACA5桁翼を891種類用意した。

またNNの構成は、教師データを変更した際の影響を調べるため、すべてのケースでFig. 1に示す同一の構造を用いた。

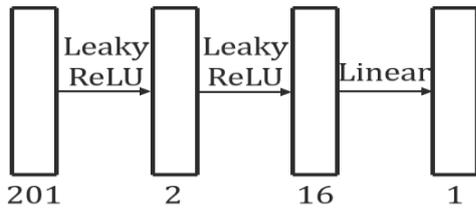


Fig. 1 Architecture of the NN used.

2つの中間層を持つ4層の全結合NNで、各層の次元は入力層で201次元（迎角1次元、形状200次元：それぞれ区別せず扱う）、出力層は1次元（非粘性計算のため揚力1次元のみ）、中間層は手前からそれぞれ2次元、16次元とした。活性化関数は出力層のみLinearで、それ以外はLeakyReLU⁽³⁾を用いた。

なお、DropoutやShortcut Connection等の構造は採用せず、入力変数の標準化のみ実施した。

3. 結果

3. 1. 物体形状の表現手法による精度の変化

物体形状の表現手法を変えて、検証を行った結果を次のTable 2に示す。

Table 2 R^2 (coefficient of determination) for each object shape-representation method.

Method	R^2
(a1) Points at equal intervals	0.997
(a2) Concentrate points on high curvature area	0.996
(a3) Fourier sine coefficients	0.991

それぞれの手法における決定係数 R^2 はほぼ同一であり、ほぼすべての検証データに対して高精度な予測ができています。ここから、形状表現を変えても、性能に大きな違いは出ないことが分かります。

(a3)の R^2 はもっとも低い結果となったが、学習プロセスが確率的操作を含むため上記の R^2 は変動しうることを、および本研究の目的である形状設計への応用を考慮すると、形状復元性の良い(a3)のFourier正弦係数を用いる手法がもっとも効果的であると考えられる。そのため以後示す結果はすべて(a3)の物体形状表現を採用した。

3. 2. 教師データ削減による影響

Table 3に、教師データを減らして学習を行った結果を示す。

迎角のパターンを等間隔に減らした(b1)と、翼型の種類を翼の番号に基づき等間隔に減らした(b2)とは、全データを利用して学習した(b0)と同レベルの R^2 を得られた。また、クラスタリングによってデータを削減した(b3)も、僅かに落ちるものの、高い R^2

を得られた。一方で、類似した翼型のみ一方で、(b4)の類似形状を集めたケースは R^2 が著しく低下し、モデルの予測精度が落ちてしまった。

Table 3 Training data reduction method and R^2

Method	R^2
(b0) 200×10^3 : All data	0.991
(b1) 25×10^3 : Angle $\Delta 2^\circ \rightarrow \Delta 16^\circ$	0.993
(b2) 25×10^3 : NACA $2n + 1 \rightarrow 16n + 1$	0.992
(b3) 25×10^3 : Clustering (k-means++)	0.979
(b4) 25×10^3 : NACA $2n + 1 \rightarrow 8749 + 2n$	0.820

またFig. 2に(b0)~(b4)の各データセットを2次元平面に射影した際のデータ分布を示す。ここで、各グラフのx軸、y軸はそれぞれ(b0)の第1主成分と第2主成分であり、図中の太実線による閉曲線は全データの凸包である。

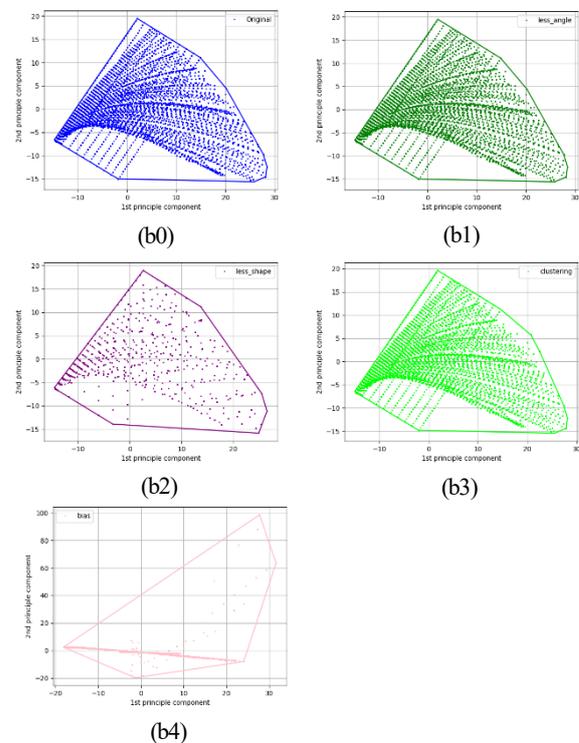


Fig. 2 Input data distribution on 2D principal components plane.

この図より、(b0)、(b1)、(b2)、(b3)の各データの分布は類似していることが分かる。また、(b4)のデータセットは他と大きく異なりまばらな分布となっていることが分かる。

これらの結果から、クラスタリングによるデータセットの削減方法(b3)は、翼型の番号や迎角など指標が無いデータセットのデータ削減時に大きな威力を発揮すると期待される。

4. 結言

本研究では、機械学習によって2次元非圧縮非粘性流の空力解析を行う手法を構築した。まず、物体形状の情報を効率良く低次元に落とし込む手法を提案し、その手法がモデル性能へ大きな影響を与えないことを明らかにした。また、教師データの多様性を維持することで、モデル性能を保ったまま教師データを削減可能なことを示した。

講演では、圧縮粘性流への拡張、設計への応用についても紹介したい。

参考文献

- (1) A. Chikitkin, M.Petrov, V.Titarev, S. Utyuzhnikov, “A multithreaded OpenMP implementation of the LU-SGS method using the multilevel decomposition of the unstructured computational mesh,” *Lobachevskii Journal of Mathematics*, 39, (2018), pp-503–512, etc.
- (2) D. Arthur, and S. Vassilvitskii, “k-means++: The Advantages of Careful Seeding,” *Proceedings of the Eighteenth Annual ACM - SIAM Symposium on Discrete Algorithms.*, (2007), pp 1027-1035.
- (3) A. L. Maas, A. Y. Hannun, and A. Y. Ng. “Rectifier Nonlinearities Improve Neural Network Acoustic Models,” *Proc. icml*, 30 (2013), pp. 3.