

第6章 流体情報学と設計情報データマイニング

6.1 流体情報学とは

これまでの流体力学は、「流れ」という通常は目に見えない運動を、定式化し測定しあるいは計算することで、目に見えるものに変えてきた。その結果は「情報」として蓄積されている。「情報」の量が絶対的に少ないときは単にその「情報」を知るだけで十分であったが、それなりに「情報」が手にはいるようになるとその「情報」をいかにうまく利用するかが重要である。そこで、実験によってもたらされる実験流体情報、数値計算によってもたらされる数値流体情報から、有益な新しい情報を創造し、新たな知の創発^{*}を促す学問として、流体情報学の構築が期待されている。本論では、流体情報学の構築に役立つと思われるいくつかの手法を概観し、将来の展望に役立てたい。

^{*}創発(emergence): 下位の階層における個々の要素の運動が上位の階層における全体的なパターンを生み、またこの上位のパターンが下位の境界条件となって個々の運動を間接的に支配すること。ここでは、下位レベルとして数値流体・実験流体の個々のシミュレーションを考えると、上位レベルにどのようなパターンが生じ下位レベルへどのようなフィードバックが生じるかを考えている。流体システムにおけるこのような創発現象を研究することでさらに多様で複雑・知的な機能を持った工学システム設計への応用が期待される。

6.2 可視化情報学

従来の数値流体でも実験流体でも可視化と呼ばれる技術が流体情報を創造する役割を担ってきたが[1]、データ量の爆発的な増加に伴い、新たな局面を迎えている。これまでのCFDでは、従来の実験で得られる限定的なデータとの照合に主眼がおかれていたため、たとえ時間・空間に複雑に分布する大量の流体データが得られてもそのほとんどを捨てていた。このため、今後の新展開を図るには、従来のデータ処理やその技法の研究でなく、大量に得られたデータをさらに統合する新たな情報技術が必要とされている。

そのためには、4次元の時空間に大量に分布する複雑な流体データを統合する、流体力学にカスタマイズされたデータマイニングの手法や、人工知能の応用などの研究が期待される。また、さまざまな物理量を時空間に加えて考えるには、さらに高次元の状態空間での可視化が必要となるであろう。

6.3 設計最適化

多数の計算・実験により得られた情報から、新たな情報を生み出す初歩的な形態の一つとして設計最適化がある。最適化では、関数評価を繰り返すことにより目的関数を最適にする設計変数の値を求める。数多くの情報から設計に有益な情報を引き出していることの例といえよう。

これをさらに進めた処理として、応答曲面法やニューラルネットワークを利用した最適化法がある。これらの手法では、本来非線形な関数の応答をうまく近似するモデルを構築することで、時間のかかるCFD計算を経ずに最適化を行う。この場合、作成された近似モデルの精度が当然結果の善し悪しを決めることになるので、精度よくかつ効率的にモデルを作成する方法が研究されている。

一般的に流れは形状の微小な変化に対して敏感であり、流体性能を最適化するために詳細形状を取り扱おうとすると、設計変数は膨大となり、最適化すべき目的関数は設計変数に対して非常に多くの多峰性を持つ関数となる。従って、勾配法で最適化を行うと、その解は初期値に強く依存する。

この限界を乗り越えるための知的な情報処理法として、進化的計算法がある[2]。進化的計算法とは、遺伝情報に基づく個体の発生と個体レベルでの繁殖淘汰をモデル化し、解集団の進化によって問題を解く、遺伝的アルゴリズム・進化戦略・進化的プログラミング等の総称である。

進化的計算法は、大域的な最適化法として知られるが、特に多目的最適化と呼ばれる分野で、トレードオフを表す解の集合であるパレート解を効率的に求めるユニークな方法として注目されている。これま

での設計トレードオフの検討が、翼厚などの特徴的なパラメータを1次元的に変化させて、流体性能の変化を調べていたのに比べ、パレート解は全設計空間における解のトレードオフをグローバルに表している。パレート解の分布を可視化することで、設計者は多くの情報を得ることができる。これは、従来の方法にない利点であるため、近年様々な分野で応用が試みられている。

6.4 データマイニング

6.4.1 データの爆発と統計学のパラダイムシフト

データ整理には、通常は統計学の知識を利用する。すなわち、生データからデータの統計量を求めて、データを遙かに少ないパラメータで記述し、母集団に関する性質(仮説)を検証する。このようにデータを縮約するための古典的な処理法として、

- 1) 母集団の分布を仮定する
- 2) 分布を決めるパラメータを推定する
- 3) 母集団に関する仮説検定を行う

という手続きを踏む。このとき、データ数(標本数)・有意水準・検定力にはある関係があり、データ数を限りなく増やすことができる場合、一定の検定力のもとでは有意水準が限りなく下がり、わずかの差を有意差として検定してしまう。逆に、有意水準を一定に保つならば検定力が限りなく1に近づき、仮説は常に棄却されてしまうことになる。

また、これらの推定・検定を確率論的に厳密に行うには標本分布を厳密に求める必要がある。この手続きは煩雑である上、一般に静的な母集団を仮定しているため、リアルタイムでデータが増えていく場合にはあまり役に立たない。しかも、データ数が増えると標本分布も正規分布に近づくので、標本分布を厳密に求めること自体の意義が薄れる。

このように、古典的な統計学の枠組みは、データを得ること自体が貴重であったときに、数少ない貴重なデータから母集団をいかに合理的に推定するか、ということを中心としていた。しかし、データがリアルタイムで大量に収集・蓄積されるようになった現在、統計学はその体系を根底から覆すパラダイムシフトに直面している。

6.4.2 データマイニング

データマイニングとは、データの爆発に対応して否応なくビジネスの現場で磨かれてきたデータ処理法の総称である[3]。使えるものは何でも使うという立場から、データマイニングの手法は古典的な多変量解析からニューラルネットまで何でも使う。しかし、データマイニングと古典的な手法の違いを際立たせているのは、なんといっても後者のニューラルネットの利用であろう。

ニューラルネットは、神経回路網のモデル化であり、機械に学習させるという点が特徴的である。この方法では、回帰分析のように特定のモデル(方程式)を仮定しない。別の見方をすると、データの主成分分析のように線形変換の積み重ねで現象を記述するのではなく、ニューラルネットという非線形関数の重み付けでデータの概要を理解しようとする。

また、古典的な統計学では、最良の推定量を採用する。すなわち、統計量の最適性・一意性を積極的に保証する。しかし、ニューラルネットでは、トポロジーやトレーニングを変えれば、得られるニューラルネットも変化する。時間をかけてベスト(最適で一意)な解を求めるより、必要なときにベターな解を得ること、データ構造をモデル化し検証するのではなく、データの構造を発見することに主眼がおかれている。

6.4.3 自己組織化マップ

自己組織化マップ(Self-Organizing Map, SOM)[4]は、近年データマイニングの一手法として脚光を浴

びている。SOM は提案者の名前を取ってコホネンネットとも呼ばれる、教師なし学習のアルゴリズムを用いるフィードフォワード型のニューラルネットモデルである。

SOM は記憶のモデル化の研究から生み出された。大脳皮質の特定の領域は、特定の感覚を受け持つように組織化されている (Fig. 6.1) [5,6]。記憶に対しても幾何学的に組織化されたマップが形成されると考えられている。この神経機能の空間的な順序づけと組織化をモデル化したものが SOM である。

SOM は多次元のデータを圧縮して、低次元 (2次元) に写像する。すなわち、SOM を用いると高次元空間の可視化ができる。SOM は、入力層と出力層の2層からなり、隠れ層はない。第2層を構成するユニットはあらかじめ平面座標を持っている (Fig. 6.2)。第1層に提示された情報はすべて第2層に提示され、第2層のユニットは入力情報にどれだけ似ているかを競争する。競争の結果選ばれた勝者ユニットは、重み付けが更新されてさらに入力に近づく。また、その近傍ユニットも近さに応じて重みを更新される。こうして隣り合うユニットは似たような重みを持つようになり、第2層に新たな位相が形作られる (Fig. 6.3)。

このように形成されたマップは、我々が直感的に理解している地図とは異なる。通常の地図では、方角 (座標軸) があり、距離 (ユークリッド距離) が定義されている。SOM には、方向性がなくユークリッド距離も定義されていない。隣り合うユニットは類似度が高いが、 n 個離れたら類似度が $1/n$ になるわけではないし、異なるユニットはまわりのユニットに対して異なる近さを持っている。その代わりとして、SOM は多次元情報を2次元に折りたたんで表現できる。すなわち、SOM を用いると多次元データの可視化ができる。本研究では SOM の作成に、Viscovery® SOMine 3.0J[7]を使用した。

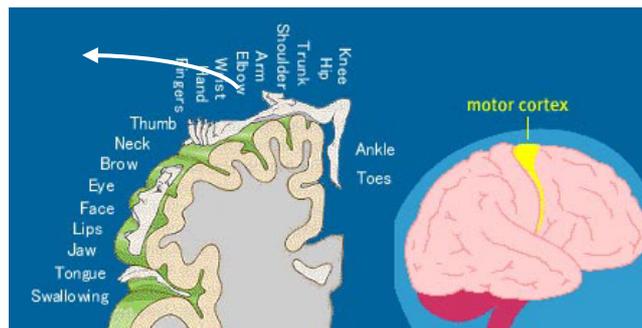


Fig. 6.1 Brain map of the motor cortex

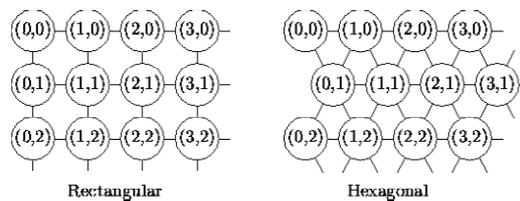


Fig. 6.2 Different topologies used in the SOM

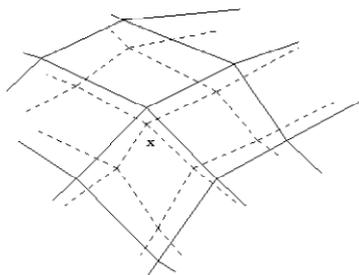


Fig. 6.3 Updating the best matching unit and its neighbors

6.5. 多目的最適化とパレート解の可視化

6.5.1 背景

筆者らは、進化的計算法の中で特に多目的遺伝的アルゴリズム(MOGA)と呼ばれる手法を、空力最適化問題に適用してきた[8-10]。特に最近の研究[10]では、4つの目的関数に対して、3次元翼を72の設計変数で表現して形状最適化を行い、766個のパレート解を求めた。これらのパレート解は、4次元目的関数空間内の3次元のトレードオフ曲面を構成している。しかし、このトレードオフの様子を把握することは困難である。目的関数が2つや3つなら図示することは明白であるが、それ以上になると高次元空間の可視化となるからである。これまでは、2次元の目的関数平面に射影したり、既存の設計と比較したりすることで、解を吟味してきたが、実際にはほんの数個の解を調べるにとどまっていた。

本研究ではデータマイニングの手法を取り入れることで、パレート面の性質を考察する。このために、コホネンの自己組織化マップを導入し、766個のパレート解によって表される3次元パレート面を可視化し、その有効性を検討する。

6.5.2 パレート解の自己組織化マップ

超音速翼の多目的最適化では、超音速巡航抵抗・遷音速巡航抵抗・翼根の曲げモーメント・ピッチングモーメントの4つを最小化した。得られたパレート解は4次元目的関数空間内の3次元曲面となる。これを2次元面で可視化する最も単純な方法は、2目的関数空間に射影することである。そこで、2目的として超音速巡航抵抗と遷音速巡航抵抗を取ると、Fig. 6.4になる。図中の曲線Iは両抵抗間のトレードオフを表す。Fig. 6.5に極限パレート解の翼平面形を示す。

これらの解から、SOMを作るとFig. 6.6のようになる。各目的関数を最小化する解を含むようなクラスタができる。またクラスタ間では、ピッチングモーメントの小さい翼と遷音速抵抗の小さい翼に類似性があり、また遷音速抵抗と超音速抵抗の小さい翼にも類似性があることが分かる。これらに共通することはアスペクト比が高いことである。図中のPareto A, Bは既存の設計より4目的すべてで優れた解であり、実用的な解はアスペクト比を抑えたものであることが確認できる。

Fig. 6.6のSOMのクラスタリングを細分化し、各領域の平均を取ることで、代表的なパレート解を48個作った。この設計変数のデータに対して、新しいSOMを作ってみた(Fig. 6.7)。これによって得られる設計変数のクラスタリングにより、パレート解を形作るためにどのような設計変数が関連を持って働いているのかが可視化できる。特に翼平面形に関する設計変数が左下にクラスタを形成しており、平面形が重要であることを裏付けている。これに続くクラスタでは、翼のキャンバーに関する設計変数が主であり、線形翼理論の指摘する重要な性質と一致している。さらに、右に移動するに連れて翼厚に関する設計変数が登場するが、これらの設計変数の重要さに関する並び具合は、一般の翼理論とよく対応する。

SOMによるクラスタリングは、パレート解の可視化のみならず、設計空間の分類・構造の解明にも役立つ。すなわち、パレート解によって与えられる大量の設計情報を仮想的な設計データベースと見なせば、SOMは設計情報データマイニングを行う優れたツールとなっている。

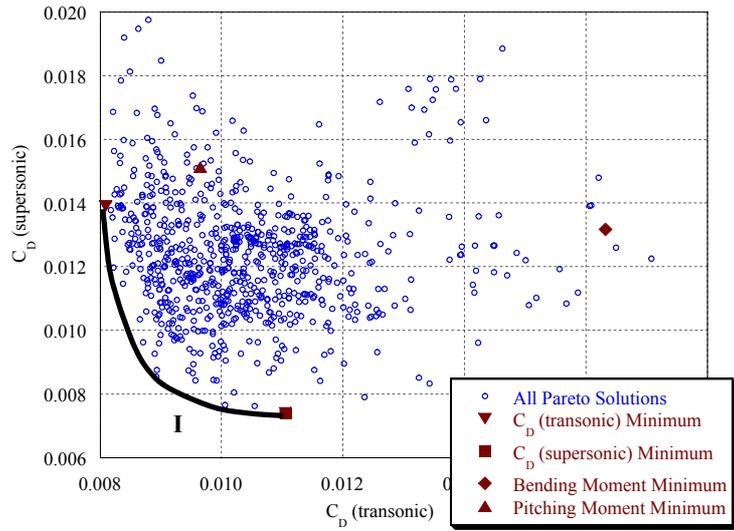


Fig. 6.4 Projection of Pareto solutions into two-dimensional plane between transonic and supersonic drag coefficients

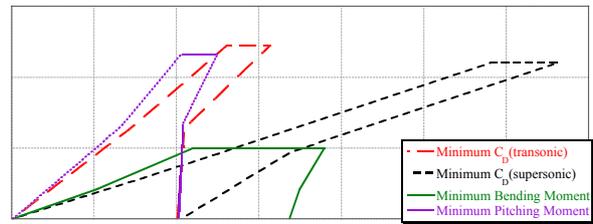


Fig. 6.5 Planform shapes of the extreme Pareto solutions

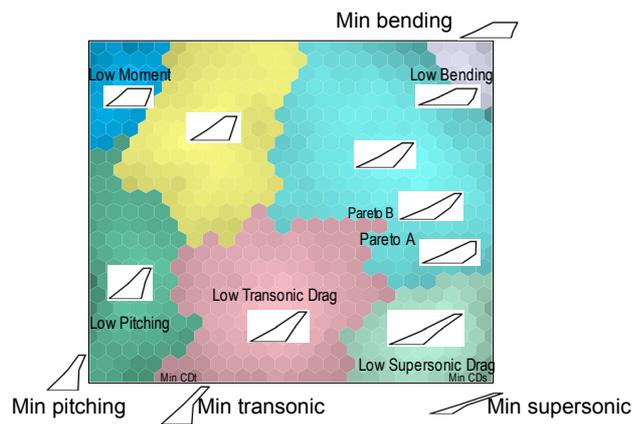


Fig. 6.6 SOM of Pareto solutions in the objective function space for supersonic wing design

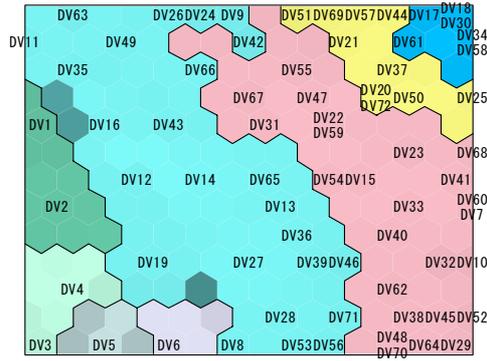


Fig. 6.7 SOM of design variable space

参考文献

- [1] 可視化情報学会入門編集委員会編, *可視化情報学入門*, 東京電気大学出版会, 東京, 1994 年.
- [2] 佐々木大輔, 竹口幸宏, 大林茂, 中橋和博, 広瀬直喜, 「次世代超音速輸送機主翼の超音速と遷音速巡航の多目的最適化」, *日本航空宇宙学会誌*, 47(551), 464-469 頁, 1999 年.
- [3] 豊田秀樹, *金鉱を掘り当てる統計学*, ブルーバックス, 講談社, 東京, 2001 年.
- [4] T.コホネン, *自己組織化マップ*, シュプリンガー・フェアラーク東京, 東京, 1996 年.
- [5] <http://www.cis.hut.fi/~jhollmen/dippa/node7.html>
- [6] <http://www.pbs.org/wgbh/aso/tryit/brain/mapcortex.html>
- [7] <http://www.eudaptics.com/technology/somine.html>
- [8] Obayashi, S., et al., "Multiobjective Evolutionary Computation for Supersonic Wing-Shape Optimization," *IEEE Transactions on Evolutionary Computation*, Vol. 4, No. 2, pp. 182-187, 2000.
- [9] Sasaki, D., et al., "Multiobjective Aerodynamic Optimization of Supersonic Wings Using Navier-Stokes Equations," *CD-Rom Proceedings of ECCOMAS2000*, Barcelona, 2000.
- [10] Sasaki, D., et al., "Navier-Stokes Optimization of Supersonic Wings with Four Design Objectives Using Evolutionary Algorithm," AIAA Paper 2001-2531, 2001.