

時空間データ分析におけるモード分解技術の活用

賀須井直規ⁱ 結城文香ⁱ 今井隆太ⁱⁱ 山出吉伸ⁱⁱⁱ

Mode Decomposition Techniques for Spatio-Temporal Data Analysis

Naoki KASUI Fumika YUKI Ryuta IMAI Yoshinobu YAMADE

製品設計や自然現象の評価の際に実験・観測・シミュレーション等を通じて取得される時空間データは大規模化・複雑化を続けているが、その効果的な活用のための分析手法として、モード分解の技術が注目を集めている。本稿では、時空間データ分析の文脈におけるモード分解の数学的概要と活用方法について概説するとともに、製造および防災分野におけるモード分解に基づく特徴抽出・次元削減・標本生成の具体的事例を通じて、その有用性を示す。

(キーワード): モード分解, 固有直交分解, 動的モード分解, 特徴抽出, 時空間データ分析

1 はじめに

製品設計や自然現象の評価に際しては、その製品や現象を特徴づける物理量に関して、実験・観測・シミュレーション等の方法により一定の空間的・時間的範囲にわたるデータを取得し分析するのが一般的である。例えば、自動車車両の設計にあたっては、走行環境を模擬した風洞実験を行って周辺流れ場を観測するし、ある地域での地震発生時の被害予測にあたっては、モデル化に基づく数値シミュレーションによって揺れの推定値を得る。

コンピュータやデータ記憶媒体の発達・低廉化、計測機器やシミュレーション技術の発展、インターネットの普及といった昨今の潮流により、こうしたデジタルデータは大規模化の一途を辿っている。一方で、大規模な時空間データはしばしば構造・傾向が複雑となり、その中から設計や評価に活用できる重要な情報を抽出することが困難になりがちである。この問題への対処法が強く期待されている。

実験・観測・シミュレーション等で得られた複雑な時空間データの構造を紐解き、本質的な情報を抽出する技術として、「モード分解」が注目されている。モード分解は、現象を構成するいくつかの特徴

的な単位成分「モード」をデータから取り出すというもので、現象の本質的な理解、簡易な推定方法の構築、新たなデータサンプルの生成等に活用できる。

本稿では、時空間データ分析に活用できるモード分解技術として代表的な2つの手法を取り上げ、その数学的概要と活用方法について記述するとともに、具体的ないくつかの活用事例を紹介することを通じて、その有用性を示す。

2 モード分解の概要

モード分解とは、データから現象を構成する特徴的な単位成分「モード」を取り出す処理のことである。本章では、モード分解技術の数学的概要と活用方法について記述する。

2.1 数学的概要

本節では、モード分解の代表的な手法である固有直交分解 (Proper Orthogonal Decomposition, POD) と動的モード分解 (Dynamic Mode Decomposition, DMD) の数学的概要について記述する。

ⁱ サイエンスソリューション部 社会インフラチーム コンサルタント

ⁱⁱ サイエンスソリューション部 社会インフラチーム 上席主任コンサルタント 博士 (数理科学)

ⁱⁱⁱ サイエンスソリューション部 社会インフラチーム 課長 博士 (工学)

2.1.1 固有直交分解 (POD)

POD は、主成分分析など時空間データ分析以外の分野においてもよく知られた特徴抽出手法と共通の技術を核とするモード分解手法である。POD では、任意の物理量の「空間分布に関する多数のサンプル」からなるデータを、全サンプルに共通する「空間構造」、空間全体に一様にかかるサンプルごとの「変動係数」、現象全体への影響度を反映する「強度」の3種の量からなる、互いに直交する「モード」の重ね合わせへと変換する。すなわち、次式(1)における u, v, σ の各添字に対応する成分を特定する。

$$\begin{cases} x_{i,j} = \sum_k u_{i,k} v_{j,k} \sigma_k \\ \sum_i u_{i,k_1} u_{i,k_2} = \delta_{k_1 k_2} \\ \sum_j v_{j,k_1} v_{j,k_2} = \delta_{k_1 k_2} \end{cases} \quad (1)$$

ここで、添字 i, j, k (および k_1, k_2) はそれぞれ観測点、サンプル、モードのインデックスであり、 $x_{i,j}$ は観測点 i ・サンプル j における元の物理量の値、 $u_{i,k}$ は k 次モードの空間構造の観測点 i における値 (無次元)、 $v_{j,k}$ は k 次モードのサンプル j に関する変動係数 (無次元)、 σ_k は k 次モードの強度 (x の次元) である。なお、 δ は Kronecker のデルタである。

具体的な手続としては、縦方向に観測点・横方向にサンプルを並べたデータ行列 X に対し、特異値分解を適用することで、 u, v, σ を成分とする行列 U, V, Σ を求めることができる。

モードは行列 X のランク (最大で、観測点数とサンプル数との小さい方) の数だけ抽出され、全モードの総和によって元のデータは完全に再現される。ただし、強度 σ に基づいて算出されるデータ全体に対する各モードの「寄与率」に注目すれば、ごく少数のモードの合計だけで 100 % 近くを占める場合が多い。このとき、特定の少数のモードのみがデータ全体を構成する情報の大部分を保持している、すなわち、それらは現象に支配的な影響を与えている重要なモードであると考えることができる。寄与率の定義は、分析対象ごとに個別に適切なものを設定する必要がある。

2.1.2 動的モード分解 (DMD)

DMD は、近年特に注目を集める時空間データ分析技術である¹⁾。DMD では、非定常現象における

任意の物理量の「時空間分布」のデータを、空間的に非一様な振幅、位相を持つ有限個の増幅/減衰振動の重ね合わせであると仮定する。この基底たる各振動成分をモードとみなし、その「空間構造特性 (振幅、位相の空間分布)」と「時間変動特性 (時間あたり増幅/減衰率、周波数)」、さらに場合によって「強度」を特定する (式(2))。

$$x_{i,j} \simeq \sum_k \sigma_k r_k^j a_{i,k} \cos(\omega_k j + \theta_{i,k}) \quad (2)$$

ここで、添字 i, k と x, σ はいずれも前述と同様の概念を指し、添字 j は時刻のインデックス、 $a_{i,k}, \theta_{i,k}$ は k 次モードの空間構造の観測点 i における初期振幅と初期位相 (いずれも無次元)、 r_k, ω_k は k 次モードの時間ステップあたりの増幅/減衰率と位相変化量 (いずれも無次元) である。

具体的な手続としては、まず、観測された物理量が、ある時刻と直前の時刻とで (平均的に) 時刻によらず一定の線形の関係にあると仮定する。すなわち、縦方向に観測点、横方向に時刻のデータ x を並べた行列 X (列数を M とする) について、それぞれ X の $1 \sim M-1$ 列目、 $2 \sim M$ 列目を取り出した行列 X_1, X_2 を用意し、近似的に次式(3)を成立させる A が存在すると考える。

$$X_2 \simeq A X_1 \quad (3)$$

次に、 A を近似した行列 $\hat{A} := X_2 X_1^+$ の固有値問題を解き、複素数として得られる固有値、固有ベクトルそれぞれの絶対値、偏角として r, ω, a, θ を求める。ここで、 X_1^+ は、 X_1 に関する Moore-Penrose の疑似逆行列である。

各モードの強度 σ を決定するための方法には種々ある。例えば、Sparsity promoting DMD では、式(2)の右辺と元データとの間の誤差に σ に関する l_1 正則化項を含めた式を σ の関数とみなして最小化することにより、ベクトル σ を疎なものとして求める。このアプローチにより、POD 同様、データを構成する最大限の情報が含まれるような少数のモード群を選択することができる。

DMD は、POD と比較して、時間発展に強い線形性・周期性がある現象においてより明確に特徴構造を抽出できる場合がある。一方で、モードの時間発展の形式に関して強い仮定があるため、抽出されたモードのいかなる重ね合わせでも元データを完全に再現することは一般にはできず、再現精度を制御することもできない点に注意する必要がある。

2.2 活用方法

モード分解で得られる各モードは、データおよび現象を構成する特徴的な単位成分に対応する。この性質から、モード分解は以下のような実用的なタスクに活用できる。

・特徴抽出：現象の挙動特性の理解・説明

モード分解は、複雑な現象に対する理解・説明に活用することができる。実験・観測・シミュレーション等で得た生の時空間データや既存の工学的分析・知見と、空間的・時間的な特徴を分離して重要度とともに抽出・可視化したモードとを複合することで、現象の本質的特性に関する分析・考察が促進される。

・次元削減：現象の情報圧縮や簡易・迅速な推定

ある条件下で得られた時空間データにおける重要モードは、しばしば類似した条件下でも同様に卓越する。例えば、ある条件下でのデータに対して POD で求めたモードの空間構造（または変動係数）を利用すれば、類似条件下で発生する現象の大部分を把握するのに、少数の重要モードの変動係数（または空間構造）のみを捕捉すればよい。このように、モード分解は、現象を記述するための未知数の削減（＝次元削減、低次元モデリング）、およびそれに基づく簡易推定手法の構築に活用することができる。

・標本生成：現象データサンプルの人工的な生成

次元削減で得られる低次元モデルは、生成モデル

としても利用できる場合がある。より具体的には、ある条件下で得られた時空間データから抽出された重要な空間構造に基づく低次元モデルを構築した上で、モデルに変動係数の組をランダムに与えることで、類似条件下での現象の新たなデータサンプル（標本）を簡便に生成することができる。多数生成した標本は、現象の統計的分析に活用することができる。

3 時空間データ分析におけるモード分解の活用例

本章では、前章末尾に示した活用方法「特徴抽出」、「次元削減」、「標本生成」それぞれについて具体的な事例を紹介することを通じて、時空間データ分析におけるモード分解技術の有用性を示す。

3.1 円柱後流の特徴抽出および次元削減

本節では、直径および主流を基準とするレイノルズ数 10^4 の円柱まわり流れを対象とした Large Eddy Simulation (LES) 結果に対し、POD を用いて特徴を抽出する事例を紹介する。LES 解析には流体解析コード FrontFlow/Blue (FFB) を使用した^{3,4)}。図 1 に、シミュレーション領域のレイアウトを示す。

まず、円柱後流の観測領域に配置した 5,151 点（x 方向に 101 点、y 方向に 51 点）の観測点における速度場を、円柱直径および主流速度で規格化した無次元時間 0.05 の間隔で 4,000 ステップ分出力した。次に、4,000 ステップ分のシミュレーションデータに対し、カルマン渦の放出周期による位相平均処理を施した。これにより、POD で抽出されるモードがより明確になることを確認している（ただし、カルマン渦の放出よりも低周波の現象に対する解析はできなくなる）。なお、カルマン渦放出周期は、円柱に作用する揚力の変動スペクトルのピークから無次元時間 5.85 ($5.85 / 0.05 = 117$ ステップ) と特定した。

本節では、上記の処理により得られるデータから速度の y 方向成分のみを取り出した 5,151（観測点数）× 117（時間ステップ数）の行列で表現されるデータに対し POD を適用してモードを抽出した。

図 2 に、分析対象データの一部を示す。

3.1.1 モード強度に基づく次元削減

抽出されたモードのうち、数値計算上の誤差を考慮して 10^{-16} 以上の固有値に対応するもののみをカウントしたところ、有効なモード数は 116 となった。

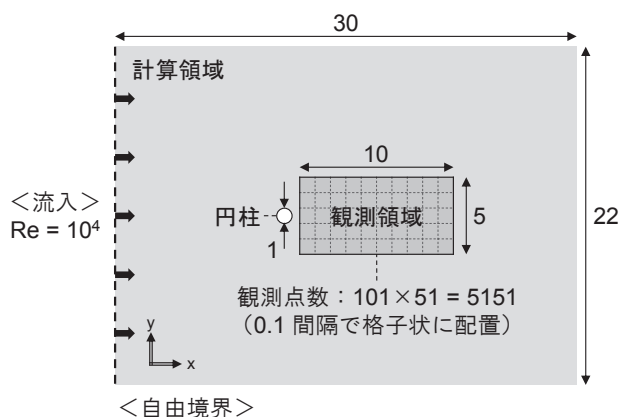


図 1 シミュレーション領域のレイアウト

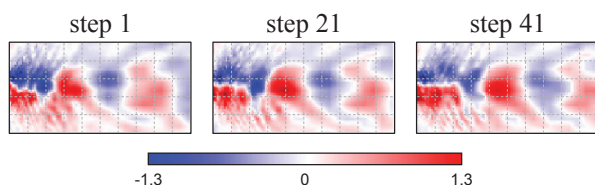


図 2 分析対象データ（無次元速度 y 方向成分）の一部

つまり、116 個のモードで現象の全体（100 %）を記述しているといえる。ここで、各モードが対応する運動によるエネルギーを反映する指標として、強度の 2 乗に基づく寄与率を計算したところ、図 3 に示すように、0 ~ 4 次の合計で 97.8 %となった（モード次数は寄与率順に与えた）。したがって、この現象は、0 ~ 4 次の 5 つのモードのみによって運動エネルギーベースで 97.8 %記述されると解釈できる。

この結果を得た上で、類似した条件下での流れ場を推定したいとき、1 ステップごとに全観測点での値を求める必要はなく、0 ~ 4 次モードの変動係数

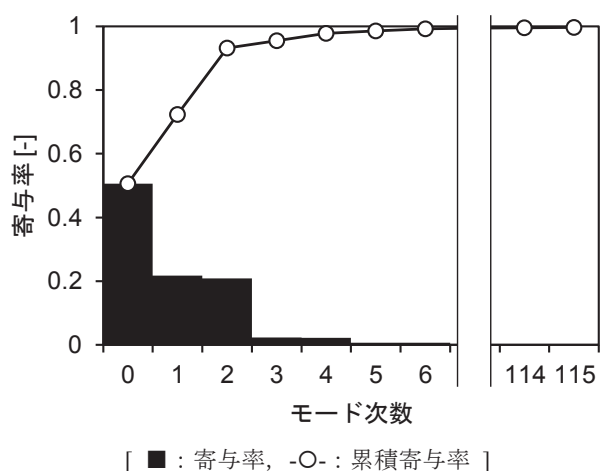


図 3 各モードの寄与率と 0 次からの累積寄与率

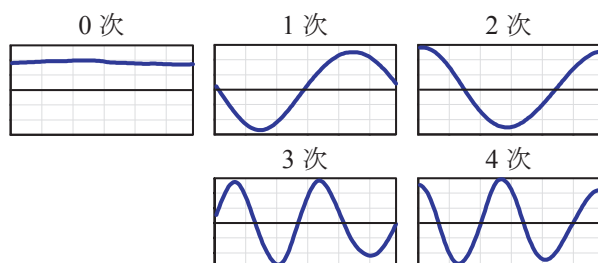


図 4 各モードの時間変動

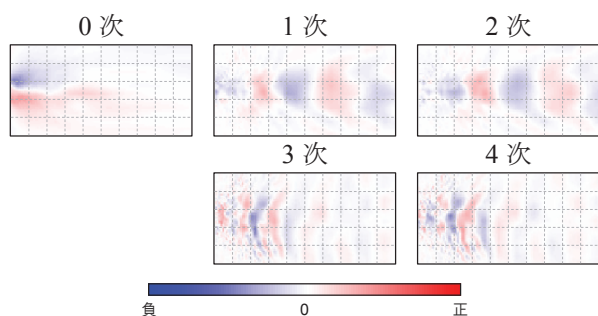


図 5 各モードの空間構造

を求めれば十分となる場合がある。その場合、1 ステップごとに決定すべき未知数の数（次元）は、5,151 から 5 に削減できる。

以上が、モード分解技術 POD を用いたデータおよび現象の次元削減の一例である。

3.1.2 モード時間変動・空間構造に基づく特徴分析

本項では、前項で重要と判定した 0 ~ 4 次の各モードについて分析する。

図 4 に、変動係数の時間変化を示す。まず、0 次モードはほとんど定常で、他のモードは 0 を中心に周期的な変動をしていることから、0 次モードは（その空間構造が）データ全体の時間平均場に対応し、他の各モードは変動場としての特定周波数での定在振動現象に対応することが分かる。

各モードの 117 ステップあたりの振動数に注目すると、変動成分のうち寄与率の大きい 1 次、2 次モードは、カルマン渦の放出現象と同じ周期を持つ現象に対応していることが分かる。このことは、モード分解により抽出されたモードが、流体力学的知見に即して考えても重要で特徴的なものであることを示唆する。3 次、4 次モードの周期は、1 次、2 次モードの 1/2 倍となっている。

また、各モードの位相に着目すると、1 次と 2 次、3 次と 4 次の各ペア内で、互いに $\pi/2$ ずれている。以上のことから、0 ~ 4 次モードは全て互いに直交するものであることが確認できる。

次に、各モードの空間構造（図 5）に注目する。0 次モードは定常場であるが、円柱近傍で流れが円柱背後に回り込む現象を捉えている。1 次、2 次モードは、円柱の直背後から細長く後方まで達するような構造を有する。時間変動の特徴と合わせて考えると、このモードはカルマン渦の放出・流下現象と強く結びついていることが示唆される。3 次、4 次モードは、1 次、2 次モードと比べて長さ（x 方向への伸び）が短く幅（y 方向の広がり）が広いほか、（x 方向への）波長が 1/2 倍となっている。

また、1 次と 2 次、3 次と 4 次の各ペア内では波長が等しく山・谷の位置が 1/4 波長分ずれており、ペア間で空間的な構造のスケールが概ね 1:1/2 となっていることは、時間変動の傾向と類似する。

最後に、各モードの強度、時間変動、空間構造の関係について考察する。周期、波長の大きさが寄与率の順位と対応しており、この結果は、抽出されたモードのうち大きな時間的・空間的構造を持つもの

が、現象全体に対してより大きな影響を及ぼしていることを示す。このこともまた、工学的な知見と整合する。

以上のように、モード分解によって、データの奥に隠れた現象の本質的特性がより明確になる。

3.2 地震動予測地図の標本生成

本節では、地震動空間分布に関する複数の標本からなるデータに関して、POD でモードを抽出しそれらを様々に合成することで、類似条件における地震動予測地図の新たな標本を、任意の数だけ簡便かつ多種多様に生成する事例について紹介する。分析対象としては、J-SHIS 想定地震地図²⁾において公開された「長野盆地西縁断層帯飯山-千曲区間」に関

するデータを使用する。これは、震源に関する 12 通りのシナリオ設定それぞれに対して地震波伝播の数値シミュレーションを行い、工学的基盤上の 18,252 箇所における最大揺れ速度を記録したデータである。

この $18,252 \times 12$ の行列で表現されるデータから POD で抽出された 12 モードの空間構造を図 6 に示す。地震動分布に対し抽出されるモードは、前節の円柱後流の場合と同様に、工学的な知見と対応付けた解釈ができる事例が報告されている³⁾。今回のモード分解結果に関して、例えば、0 次モードは平均的な距離減衰特性、2 次モードは地盤特性、3, 4 次モードはディレクティビティ効果を反映したモードである、といった解釈ができる可能性がある。た

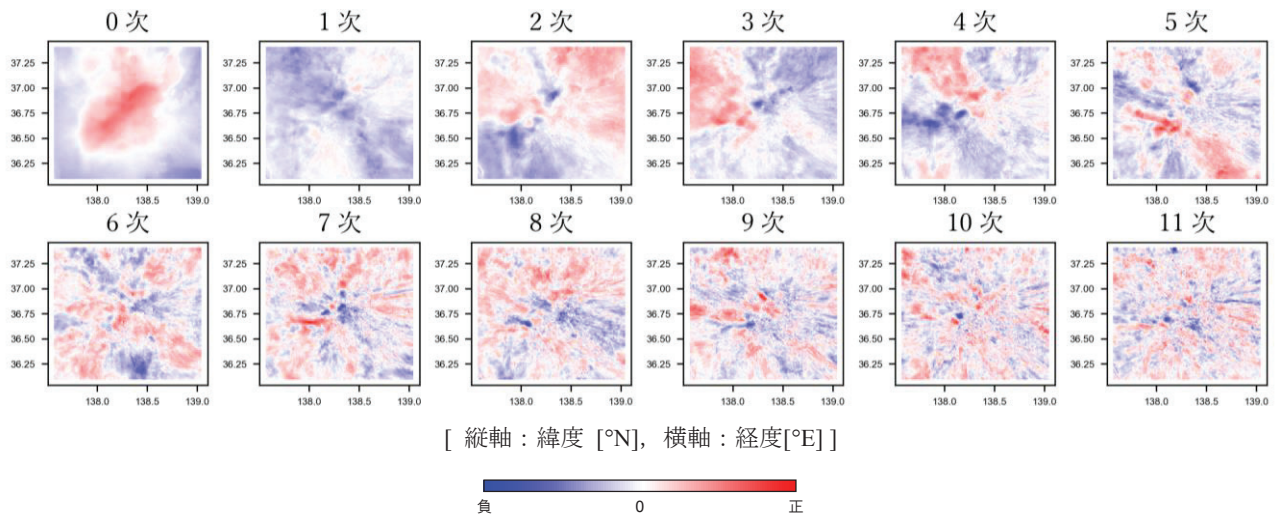
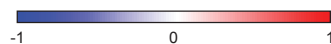


図 6 各モードの空間構造

表 1 各標本の各モードに対する変動係数

	0 次	1 次	2 次	3 次	4 次	5 次	6 次	7 次	8 次	9 次	10 次	11 次
標本 1	0.305	0.245	0.279	0.150	0.026	0.352	0.084	0.019	0.123	0.429	0.644	0.034
標本 2	0.278	0.242	-0.048	-0.100	-0.439	-0.144	-0.253	-0.529	-0.054	-0.117	0.026	0.527
標本 3	0.293	0.138	-0.255	-0.249	-0.031	-0.192	0.022	0.329	-0.220	0.667	-0.337	0.117
標本 4	0.277	0.090	-0.377	0.295	0.340	-0.196	0.222	-0.451	-0.395	-0.016	0.110	-0.332
標本 5	0.313	0.377	0.394	-0.150	0.307	0.401	0.111	-0.013	-0.171	-0.279	-0.456	0.019
標本 6	0.286	0.132	0.197	0.085	-0.263	-0.277	-0.017	-0.153	0.561	0.087	-0.259	-0.543
標本 7	0.294	0.213	-0.130	-0.400	-0.101	-0.290	0.125	0.432	-0.043	-0.470	0.391	-0.138
標本 8	0.289	-0.054	-0.234	0.265	0.511	-0.144	-0.099	0.155	0.524	-0.124	-0.050	0.425
標本 9	0.280	-0.290	0.212	0.552	-0.378	-0.085	0.385	0.273	-0.222	-0.130	-0.119	0.184
標本 10	0.269	-0.140	-0.524	0.086	-0.276	0.621	-0.290	0.114	0.045	-0.124	-0.075	-0.212
標本 11	0.294	-0.381	0.355	0.001	0.178	-0.179	-0.680	0.066	-0.290	0.006	0.088	-0.142
標本 12	0.284	-0.630	0.038	-0.496	0.062	0.117	0.384	-0.285	0.146	0.053	0.030	0.056



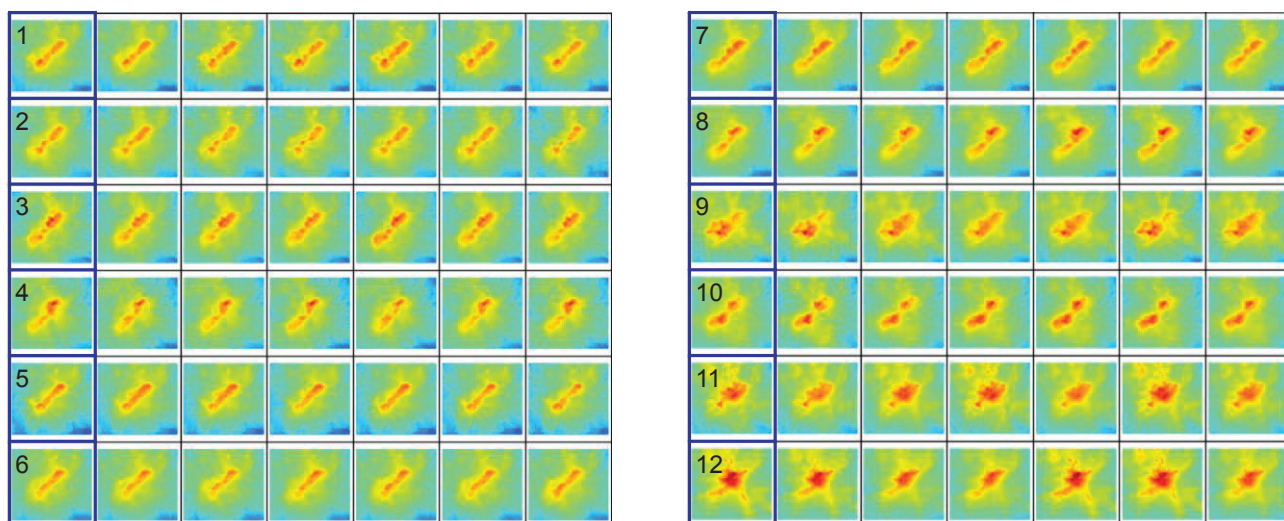


図7 既存12標本と生成された72標本

* 番号付きのものは既存標本

だし、工学的知見とモードとの対応付けに関する詳細な議論には、計算条件に関する大規模で複雑なモデルデータが必要となるため、ここでは省略する。

次に、空間構造を保存したまま、シナリオを特徴づける変動係数として新たなものを与えることで、新たなシナリオに対応する地震動予測地図の標本を生成する。このとき、既存の標本における変動係数（表1）の分布傾向に従うような変動係数の組を設定することが重要である。具体的な設定方法については、紙面の都合上ここでは省略する。

標本生成の実行例として、既存の12標本から新たに72標本を生成した（図7）。生成された標本は、既存の標本群と共通した分布傾向を示していると同時に、どの標本とも異なっていることが確認できる。このことから、PODを用いて、既存データの特徴を捉えた新たな標本を生成することができたといえる。

このような標本生成は、多数の地震動予測地図を用いたハザード評価の高速化において有用であることが期待できる。地震動予測地図の標本を得るためには、通常、地盤物性分布や震源断層特性等の複雑な条件を手動設定して大規模なシミュレーションを行うことになるが、このプロセスには、スーパーコンピュータを用いても数十時間を要する場合がある。PODを用いれば、このような標本が既に蓄積されている状況であれば、既存データの特徴を抽出して生成モデルを構築し、一般的なノートPCでも数秒程度で新たな標本を生成できるようになる。

4 おわりに

本稿では、実験・観測・シミュレーション等で得られるデータの大規模化・複雑化という昨今の潮流を踏まえ、時空間データ分析の文脈におけるモード分解の技術について、その数学的概要と活用方法について記述するとともに、具体的ないくつかの活用事例を紹介した。

モード分解の代表的な活用方法として、特徴抽出、次元削減、標本生成の3つがある。円柱後流データを例に、モード強度に基づく次元削減と、抽出された変動係数と空間構造とに基づく挙動特性の分析を通じ、時空間データの特徴抽出および次元削減におけるモード分解の有用性を示した。また、地震動空間分布を例に、既存データに見られる特徴的な傾向を持つ新たな標本の生成において、モード分解を活用できることを具体的に示した。

本稿ではモード分解の手法として特にPODに焦点を当てて具体例を示したが、その他にも、2章に示したDMDや畳み込みニューラルネットワーク（Convolutional Neural Network, CNN）を活用した手法等、同様の目的に活用できる技術に関する研究・開発が加速している。今後、モード分解の分野が更なる発展を遂げ、様々な形式・性質をもつ時空間データを、目的に応じより適切に分析できるようになることが期待される。

当部では、主柱業務である科学技術計算サービスに加え、こうしたデータ分析サービスにも今後ますます注力し、研究開発の技術的支援をより多角的に実施していく予定である。

謝辞：本稿で使用した J-SHIS 想定地震地図データは国立研究開発法人防災科学技術研究所からご提供いただいたものである。ここに深い謝意を表す。

引用文献

- 1) 例えば, Jonathan H. Tu, Clarence W. Rowley, Dirk M. Luchtenburg, Steven L. Brunton, J. Nathan Kutz: On Dynamic Mode Decomposition: Theory and Applications, J.Comput. Dyn. 1(2), (2014) 391-421.
- 2) <http://www.j-shis.bosai.go.jp/map/>, Accessed on 2020.7.31.
- 3) C. Kato, M. Kaiho, and A. Manabe: An overset finite-element large-eddy-simulation method with application to turbomachinery and aeroacoustics, Transactions of ASME, Journal of Applied Mechanics, vol. 70, (2003) 32-43.
- 4) C. Kato, Y. Yamade, H. Wang, Y. Guo, M. Miyazawa, T. Takaishi, and Y. Takano: Numerical prediction of sound generated from flows with a low Mach number, Computers & Fluids, vol. 36, no. 1, (2005) 53-68.
- 5) 能島暢呂, 久世益充, LE QUANG DUC: シナリオ地震動予測地図の特異値分解によるモード分解と地震動分布のシミュレーション, 日本地震工学会論文集 第18巻 第2号, (2018) 95-114.