

機関番号： 14501

研究種目：若手研究(B)

研究期間：2009～2010

課題番号：21740344

研究課題名(和文) 非線形次元削減法を用いた非線形システムにおける確率分布およびレジームの検出

研究課題名(英文) Detection of probability distribution and regimes in nonlinear systems with nonlinear dimension reduction methods

研究代表者 西澤 誠也 (NISHIZAWA SEIYA)

神戸大学・大学院理学研究科・助教

研究者番号：40447892

研究成果の概要(和文)： 大気システムは、超高次元の非線形システムであり、高次元のままではその変動の解析は困難である。本研究では、冬季極域成層圏気温変動や降水変動といった非線形性が強い大気の変動から、非線形次元削減法を用いて低次元の多様体の抽出を行った。線形次元削減法を用いた場合には原理的に正しく表現できない物理変数の正負の非対称性を、非線形手法を用いることで正しく反映した多様体を得ることができた。また、その多様体上において、状態変数の存在確率分布やクラスターの推定を行い、有意では無いもののレジームの存在を示唆する結果が得られた。

研究成果の概要(英文)： The atmospheric system is a high dimensional nonlinear system and naïve analyses of variability in the system is not easy due to the high dimensionality. In this study, I extract low dimensional manifolds from strongly nonlinear atmospheric variability, such as temperature in winter polar stratosphere and precipitation, with nonlinear dimension reduction methods. Those manifolds show asymmetry in physical quantities between positive and negative values, although the asymmetry cannot be represented with linear methods. Additionally I estimate existence probability distribution of state variables and clusters on the manifolds, and obtain possible regimes, though it is not statistically significant.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009年度	2,000,000	600,000	2,600,000
2010年度	1,200,000	360,000	1,560,000
年度			
年度			
年度			
総計	3,200,000	960,000	4,160,000

研究分野：気象力学

科研費の分科・細目：気象・海洋物理・陸水学

キーワード：非線形次元削減

1. 研究開始当初の背景

あるシステムにおける変動の状態は、気温などの各物理変数の空間分布の組み合わせとして定義される。各物理変数の実空間における各格子点上の値を状態変数としたとき、状態変数を軸とする多次元空間を状態空間と呼ぶ。ある瞬間の大気変動の状態は、状態空間における1点(ベクトル)に対応し、変動の時間発展は、トラジェクトリに対応づけられる。一般に、気候システムの状態空間は、非常に高い次元をもつ空間であり、高次元な状態空間上でものごとを考えることは困難である場合が多い。したがって、低次元な空間上考えるために、次元削減(dimensionality reduction)を行う。次元削減により高次元状態空間のなかに埋め込まれている低次元多様体(manifold)を抽出することにより、変動の重要な性質は残したまま、解析対象の空間次元を下げる。そして、その多様体上における状態変数の確率分布を考えることにより、システムが状態空間上で好む場所(レジーム)を知ることができ、システムの変動の性質を理解することが可能である。

2. 研究の目的

本研究では、非線形システム、特に気候システム、を取り上げ、その状態空間の中に埋め込まれている多様体上を抽出し、その多様体上での確率分布を求める。そして、その非線形なシステムのレジームを検出し、非線形変動の性質の理解に結びつけることを目的とする。大気変動の研究では、変動の線形性および正規性を仮定しているものが多い。これらの仮定により、多くの強力な解析手法を適用することができ、多くの重要な結果が得られている。しかし、これら線形性や正規性の仮定により、実際の変動の性質のうち見えなくなってしまうものが存在したり、謝った解釈を導いたりする可能性がある(Nishizawa and Yoden 2005 など)。線形システムにおいて、次元削減のために用いられる方法として、主成分分析(Principal Component Analysis; PCA)、独立成分分析(Independent Component Analysis; ICA)や多次元尺度構成法(Multidimensional Scaling; MDS)などがあげられる。PCAは、気象・海洋学分野では経験直交関数(Empirical Orthogonal Function; EOF)解析と呼ばれることが多い。これらは、線形システムにおいては、非常に有用な次元削減方法である(Storch and Zwiers 1999 など)。線形な方法によって抽出可能な多様体は、ユークリッド空間のみである。非線形性が強い変動の場合、多様体は複雑な構造をしている場合が多い。このような変動に対し、線形の方法によりその多様体の抽出をしようとす

ると、多様体を包含するユークリッド空間を抽出する必要があり、多様体の次元数よりも多くの次元を必要とし、かつ余分な空間まで抽出してしまう。また、線形的手法によって非線形の変動を抽出することはでき無いため、得られた個々の主要変動は、実際とは異なったゆがめられた変動である可能性がある。

そこで、本研究では、非線形次元削減法を利用し、多様体の抽出を試みる。状態空間における変動の全分散のうち、抽出された低次元空間における変動によって説明される割合は、非線形次元削減法を利用した場合、一般に、同じ次元数の線形次元削減法の場合と比べて高くなる。また、低次元空間の次元数に対する分散の割合の依存性から、埋め込まれている多様体の次元数を見積もることができるが、一般に、非線形手法の場合の方が線形手法の場合と比較して正確に見積もることができる(Tenenbaum et al. 2000, Agrafiotis and Xu 2002, Gamez et al. 2004 など)。

ただし、非線形次元削減法を利用する際にもいくつかの注意点がある。非線形次元削減を行うための手法は複数あり、かつ、それぞれの手法において、複数の任意パラメータが必要である場合が多い。そして、それぞれの手法やパラメータによって異なる結果が得られる可能性がある。したがって、非線形次元削減を行う上では、複数の手法やパラメータによる結果を比較し、ロバストな結果であるか確認する必要がある。また、非線形次元削減法を利用する上で、大きな問題として、過剰適合(over fitting)が挙げられる。いくつかの手法においては、回避する方法の提案がなされている(Finnoff et al. 1993, Christiansen 2005 など)。

気候データに対し、非線形次元削減法を適用している研究は多くない(Monahan 2001, 2002, Hsieh 2001, 2007, Monahan 2002, Gamez et al. 2004 など)。また、これらの研究は、それぞれ1種類のみ手法しか用いられていないため、その結果は、用いられている手法に固有な見かけのものである可能性がある。したがって、本研究では、複数の非線形次元削減法を用いて考察を行い、手法に依らないロバストな結果を得ることを試みる。

気候システムの状態空間における確率分布は、非正規性をもっていることが知られている。特に、状態空間中に、クラスタが存在する領域があることが指摘されており、これらの領域は、気候レジームと呼ばれる(Corti et al. 1999, Parmer 1999, Stephenson et al. 2004 など)。これらレジームは、カオスシステムにおけるアトラクタと結びつけて議論されることが多い(Corti et al. 1999, Parmer 1999 など)。レジームの検出には、クラスタ解析

(Mo and Gill 1988, Cheng and Wallace 1993 など)や状態変数の同時確率密度関数(Probability Density Function; p.d.f.)の多モード抽出が用いられる(Sutera 1986, Hansen and Sutera 1986, Molteni et al. 1988, 1990, Kimoto and Ghil 1993, Corti et al. 1999 など)。ただし、これらのレジームの存在の統計的有意性に関しては、さまざまな立場の研究がある(Corti et al. 1999, Christiansen 2002, Stephenson et al. 2004 など)。

次元削減に線形的手法を用いた場合、次元削減された空間上のある 1 つの軸に対し、複数の変動成分が対応づけられていることがあるため、中心極限定理により得られる分布は実際の分布よりも正規分布に近い可能性がある。そのため、レジームの検出が困難となっている可能性があり、レジームの検出には非線形次元削減法が効果的である考えられる。非線形次元削減された空間上で確率分布を考えている研究は私が知る限りは Monahan (2001, 2002), Christiansen(2005) だけである。Monahan(2001,2002)は、five-layer feed-forward neural network 法を用いて、1 次元空間上で分布関数を求め、レジームの検出を行った。しかし、Christiansen(2005)は、Monahan(2001,2002)が用いている方法では、元の変動が正規な場合でも多モードの分布が得られることを示し、彼らの結果には問題があることを指摘している。本研究では、気候データを用いて、複数の非線形次元削減法により得られた 2 次元以上多面体上における分布関数を考え、レジームの検出を試みる。

3. 研究の方法

(1) 複数の非線形次元削減法のプログラミング実装を行い、空間に埋め込まれている多様体の性質が分かっているデータに対して、それらを適用し、それぞれの手法の特徴を比較・考察する。

(2) また、それらを非線形性が高いと考えられる気候データに適用し、気候システムに埋め込まれている多様体を抽出し、線形手法によって抽出されたものと比較し、考察を行う。

(3) 抽出された多様体上で存在確率分布を見積もり、その非正規性の検定を行う。非正規性が統計的に有意な場合は、モードの抽出およびクラスタ抽出により、気候レジームの検出を行う。

4. 研究成果

(1) 複数の非線形次元削減法を理想化データ、および現実大気データに適用し、検証および、それぞれの手法の特徴の考察を行った。局所線形埋込(Locally linear embedding, LLE; Roweis and Saul 2000), Isomap(Tenenbaum et al. 2000), Laplacian Eigenmap(Belkin and Niyogi 2003), および線形次元削減法として PCA それぞれの解析プログラムの実装を行った。そして、多様体を埋め込んだ人工的な低次元理想化データを用いてそれぞれの手法の特徴を調べた。また、それぞれの手法には、任意パラメータを与える必要があるが、期待された結果が得られるためのパラメータの選び方の検証を行った。これらのパラメータは、問題によって適切な値が異なるため、それぞれの問題で、パラメータを変えて計算し、結果の尤もらしさと、パラメータの値の変化に対する敏感性が小さいことをもって値の決定を行った。

(2) 次に、それぞれの手法を冬季極域成層圏気温データに適用し、極夜ジェット振動の抽出を行った。極夜ジェット振動は、成層圏突然昇温現象および極夜ジェットによる極渦隔離による放射冷却というように、正偏差と負偏差の原因となる物理過程が異なっている。したがって、線形手法ではうまく抽出できない可能性がある。それぞれの手法で抽出された変動の違いを比較し、それらの違いの考察を行った。その結果、線形および非線形のすべての手法で極夜ジェット振動を抽出することができた。しかしながら、それぞれ抽出された多様体の構造や、多様体上の時間発展の特徴には違いがみられた。特に、非線形手法では、変動の正負の偏差に対応する領域での特徴に違いがみられることが分かった。

(3) また、非線形次元削減法を日本域の格子点降水量データ (APHRODITE: <http://www.chikyu.ac.jp/precip/>) に適用した。降雨データは負の値はとらず、かつデータの多くが 0 付近に集中する特徴を持っている。したがって、従来の線形的な手法を用いることは不相当である可能性がある。線形次元削減法と、非線形手法を適用し、それらの結果を比較した。線形的手法である主成分分析では、第 1 モードに日本全域にまたがる構造、第 2 モードに東・西日本で逆符号をもつ構造といった空間的に大きな構造が得られた。一方、非線形手法では、沖縄地域や東海地方に集中した構造など、より局所的な構造が抽出された。現実の降雨現象が高い局所性を持つという特徴を考えると、非線形手法の方が現実の現象をうまく抽出していると考えられる。したがって、非線形次元削減の有用性を示すことができた。非線形性が高い

と考えられる他の現象においても、非線形次元削減法が有効であると考えられる。また、抽出された低次元空間において、存在確率分布を推定するとともに、クラスタの検出を行った。確率分布の推定には Kernel Density Estimation 法を用いた。また、クラスタの推定には、k-means 法を用いた。ランダムな擾乱(レッドノイズ)の成分が大きいため、シグナルとノイズをうまく分離することが難しく、有意なレジームを検出することはできなかったが、レジームの存在を示唆する結果が得られた。今後、複数の視点からレジームの存在を検証し、それらの力学的解釈を行う予定である。

(4) また、本研究では、多くの解析手法やデータを組み合わせて使用するため、データや解析プログラムの種類が多く、注意深く管理する必要がある。そこで、それらをデータベースにより管理するための Web ベースのサーバソフトウェアを開発し、本研究の効率的な推進のために利用した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- ① Nishizawa, S., T. Horinouchi, 他 5 (1 番目), Gfdnavi, web-based data and knowledge server software for geophysical fluid sciences, Part II: RESTful web services and object-oriented programming interface, LNCS, 査読有, Vol.6193, 2010, pp.105-116
- ② Horinouchi, T., S. Nishizawa, 他 6 (2 番目), Gfdnavi, web-based data and knowledge server software for geophysical fluid sciences, Part I: Rationales, stand-alone features, and supporting knowledge documentation linked to data, LNCS, 査読有, Vol.6193, 2010, pp.93-104
- ③ Kohma, M., S. Nishizawa, S. Yoden, Classification of polar-night jet oscillations and their relationship to fast and slow variations in a global mechanistic circulation model of the stratosphere and troposphere. Journal of Climate, 査読有, Vol. 23, 2010, doi:10.1175/2010JCLI2458.1

[学会発表] (計 1 件)

- ① Nishizawa, S. et al. Gfdnavi, web-based data and knowledge server software for geophysical fluid sciences,

Part II: RESTful web services and object-oriented programming interface, DIEW 2001, 2010/4/4, 筑波大学 (茨城県)

6. 研究組織

(1) 研究代表者

西澤 誠也 (NISHIZAWA SEIYA)

神戸大学・大学院理学研究科・助教

研究者番号: 40447892