

オートエンコーダを用いた次元削減による宇宙機の異常検知

Dimensionality Reduction with the Autoencoder for Anomaly Detection of Spacecrafts

櫻田麻由*1 矢入健久*2
Mayu Sakurada Takehisa Yairi

*1 東京大学大学院 工学系研究科 航空宇宙工学専攻
Department of Aeronautics and Astronautics, The University of Tokyo

*2 東京大学 先端科学技術研究センター
RCAST, The University of Tokyo

In this research, we propose to use the autoencoder in dimensionality reduction for anomaly detection of spacecrafts. We show that the autoencoder is useful compared to two conventional methods: linear PCA and kernel PCA. The autoencoder can improve the accuracy of anomaly detection by over-complete set of basis. In addition, it can perform non-linear encoding faster than conventional approaches, and also avoid over-fitting problem.

1. はじめに

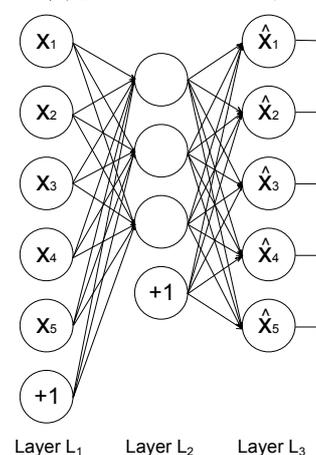
人工衛星やロケットをはじめとする宇宙機は、複雑なシステムを有し、かつ高い信頼性を求められる。一度宇宙に打ち上げると直接点検をすることができないため、地上局に送信されるデータのみから宇宙機の状態を診断しなければならない。そのため、宇宙システムにおいては、データから異常を素早く正確に発見することが求められる。こうした課題に対して、機械学習やデータマイニング手法を用いたデータ駆動型の異常検知手法が過去にも先行研究 [矢入 11] などにおいて提案されており、次元削減およびクラスタリングを宇宙機テレメトリデータに適用し、異常の検知に成功している。先行研究 [矢入 11] では次元削減手法として主成分分析 (PCA)、特に線形 PCA とカーネル PCA が用いられている。線形 PCA は最も基本的な線形次元削減手法であるが、変数同士の相関が非線形である場合には次元削減が不十分となってしまう、正しく異常検知が行えない。一方カーネル PCA は代表的な非線形次元削減手法であるが、カーネル行列の固有値算出などで計算コストが高い点 [Hoffman 07] や、次元削減を特徴空間上で行うことから入力と再構成後の出力を簡単に比べられないという点で実用上の難点がある。そこで本研究では、入力と同空間において非線形の次元削減が行えるオートエンコーダを用いて、宇宙機テレメトリデータの異常検知を行い、線形 PCA、カーネル PCA と比較することでその有用性を示した。

2. オートエンコーダによる異常検知

次元削減による異常検知では、正常データと異常データがそれぞれ別の領域に分布するような部分空間に転写することが可能であるという前提に基づいている [Chandola 09]。宇宙機のテレメトリデータは数十から時に数千に及ぶ様々な入力を持つが、一般にこれらの入力は互いに相関を持つため、冗長な入力を除き、データを低次元のベクトルに落とす表現することが可能である。

D の入力変数を持った各データサンプルを $x^{(i)} \in \mathbb{R}^D$ とし、 m 個のデータから成るデータセットを $\{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$ とする。オートエンコーダは教師無しのニューラルネットワークの一つであり、入力を再現するような出力 $\{\hat{x}^{(1)}, \hat{x}^{(2)}, \dots, \hat{x}^{(m)}\}$ を学習によって得ることが目的である。図 1 にオートエンコー

図 1: オートエンコーダ

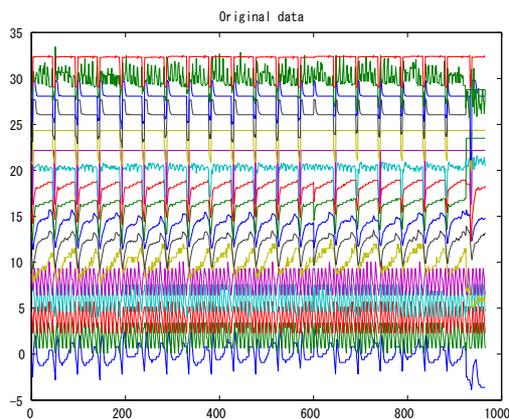


ダの概念図を示す。この図に示したように、出力層と入力層は同じ数のニューロンを持っており、出力層では入力層を再現するように同じ次元のベクトルを出力する。通常隠れ層ではニューロン数が入力層よりも少なくなっており、この層において入力は圧縮され、次元削減が行われる。 l 番目の層の i 番目のユニットの活性化度は、式 1 で表される。

$$a_i^{(l)} = f \left(\sum_{j=1}^n W_{ij}^{(l-1)} a_j^{(l-1)} + b_i^{(l)} \right) \quad (1)$$

ここで、 W と b はそれぞれ結合荷重とバイアスのパラメータである。図 1 の例では、入力層では $a^{(1)} = x$ であり、出力層では $a^{(3)} = \hat{x}$ が成り立つ。 f は活性化関数であり、本研究では隠れ層への活性化度の変換にシグモイド関数 $f(x) = 1/(1+e^{-x})$ を用いている。ただし出力層においては、今回は $[-1, 1]$ の範囲外にも値を取りうるため線形関数 $f(x) = x$ を用いている。オートエンコーダの学習、つまり訓練は W と b を決定することと同値であるが、これは勾配法などの反復解法によって式 2 に示す目的関数を最小にする W , b を求めることで決定さ

図 2: 正規化処理済み元データ



れる。

$$J(W, b) = \frac{1}{m} \sum_{i=1}^m \left(\frac{1}{2} \|x^{(i)} - \hat{x}^{(i)}\|^2 \right) + \frac{\lambda}{2} \sum_{l=1}^{n_l-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (W_{ji}^{(l)})^2 \quad (2)$$

異常の判定には、式 3 で表される再構成誤差を用いる。これはそれぞれの訓練・試験データに対して入力と出力の平均二乗誤差をとったものである。正常なデータによってモデルの学習がなされることで、正常な試験データに対しては正しく再構成が行われ誤差は小さくなる一方、異常なデータに対しては学習がされていないため再構成誤差が大きくなる。

$$Err = \sqrt{\sum_{j=1}^D (x_j^{(i)} - \hat{x}_j^{(i)})^2} \quad (3)$$

3. 実験

3.1 概要

対象のデータには、先行研究 [矢入 11] で用いられたものと同様の、つばさと呼ばれる民生部品・コンポーネント実証衛星 MDS-1 のデータを用いた。図 2 に使用したデータを示す。次元削減手法を適用する前に、入力を線形変換により平均を 0、分散を 1 に正規化して扱った。2003 年 7 月 29 日から 30 日の間の全 965 ステップの内、600 ステップまでを訓練データ、後の 365 ステップを試験データとし、各ステップにおける 17 変数の値を入力とした。本衛星において、試験データに含まれる 960 ステップ目付近で実際に異常が発生したことがわかっている。また、今回扱った 17 の変数は、機器温度、電圧、電流等の連続値である。

線形 PCA、カーネル PCA、オートエンコーダの 3 つの次元削減手法について、それぞれ実験を行った。式 2 に示したオートエンコーダの目的関数において、正規化項パラメータは $\lambda = 0.001$ に固定した。次元削減を行った後、それぞれの手法に対して式 3 で表される再構成誤差を比較した。合わせて、再構成されたデータと元のデータとの差分についても比較を行った。

3.2 結果

図 3 によると、潜在次元を 17 次元に設定した線形 PCA 以外の手法では、930 ステップ付近での再構成誤差が有意な差を持って大きくなっており、この点付近で生じた異常をオートエンコーダによる次元削減によって検知できていることがわかる。

図 3, 4 の (a)(b)(e)(f) において線形 PCA とオートエンコーダを比較してみると、潜在次元を 6 次元とした場合には両者に差はないものの、潜在次元をさらに増やし 17 次元としたとき、オートエンコーダでは正常時の再構成誤差をより小さくして異常時との差を顕著にさせることができ、再構成をより精度よく行うことができたと言えるが、線形 PCA では正常時と異常時との差が小さくなってしまっている。これは、線形 PCA では潜在次元数が入力次元と同数であることで次元削減に失敗している一方、オートエンコーダでは潜在次元を増やした場合でも過完備な基底を形成し次元削減ができるためであると考えられる。

またカーネル PCA とオートエンコーダを比較してみると、図 3(c)(d) からカーネル PCA では潜在次元を増やすことでオートエンコーダと同じように正常時の再構成誤差を小さくすることができていることがわかる。しかし図 4(d) を見ると、訓練データ期間である 1 から 600 ステップまでは再構成後データと元データの差は小さいが、試験データ期間の 600 ステップ以降から突然差異が大きくなっているのがわかる。これは、カーネル PCA の訓練において高次元のカーネル特徴空間に入力データを非線形写像することで、訓練データへの過学習が起こりやすくなっていることによるものと考えられる。一方オートエンコーダでは、潜在次元をどれだけ増やしても正規化項の効果によって訓練データに過学習を起こすことはない。

4. 結論

本研究では、宇宙機の異常検知においてオートエンコーダを用いた次元削減による異常検知が有用であることを示した。その根拠の一つには、線形 PCA と比較した場合に、潜在次元を増やして過完備な基底を形成することでより精度のよい次元削減による異常検知が可能になることが挙げられる。もう一つには、カーネル PCA と比較した場合に、計算コストを抑えた非線形変換を行うことができ、再構成された出力を入力と同空間上で簡単に比べられることや、訓練データへの過学習を避けられることが挙げられる。

参考文献

- [Bengio 09] Bengio, Y.: Learning Deep Architectures for AI, in *FTML* (2009)
- [Chandola 09] Chandola, V., Banerjee, A. and Kumar, V.: Anomaly Detection: A Survey, in *ACM Computing Surveys*, Vol. 41(3), Article 15 (2009)
- [Hoffman 07] Hoffman, H.: Kernel PCA for novelty detection, in *Pattern Recognition*. Vol. 40, Issue 3, pp.863-874 (2007)
- [Hinton 06] Hinton, G. E. and Salakhutdinov, R. R.: Reducing the dimensionality of data with neural networks, in *Science* (2006)
- [矢入 11] 矢入健久, 乾稔, 河原吉伸, 高田昇: 次元削減とクラスタリングによる宇宙機テレメトリ監視法, 日本航空宇宙学会論文集, Vol. 59, No. 691, pp.197-205 (2011)

図 3: 再構成誤差

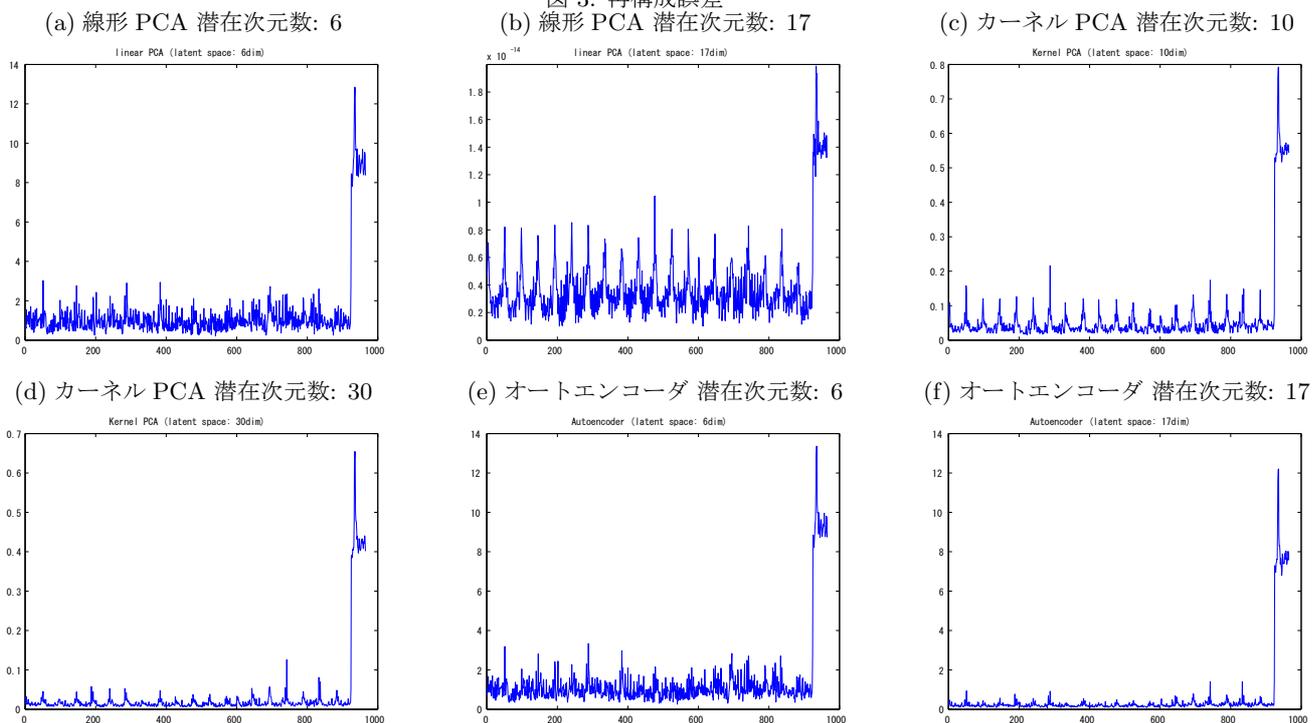


図 4: 再構成データと元データの差異

