

オンライン状態空間学習による安定性監視法

Stability Tracking by Online State-Space Learning

河原吉伸*1

Yoshinobu KAWAHARA

矢入健久*2

Takehisa YAIRI

町田和雄*2

Kazuo MACHIDA

*1 東京大学大学院 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

Research Center for Advanced Science and Technology, The University of Tokyo

Stability is one of the most important attribute for dynamical systems. In this paper, we propose a new framework for capturing the changes of dynamics behind available input-output time-series data, based on the comparison from the aspect of stability. As metrics for such purpose, we define the principal angles and distances between the subspaces spanned by the observability matrices of state-space models, which can be expressed using the poles of the models. Using the metrics, the change of stability can be estimated recursively by learning a state-space model from input-output time-series data on real-time. The application of the proposed framework to the fault-detection problem is described, and the experiment using an artificial data shows the usefulness.

1. 序論

動的システムの安定性は、産業プラントや宇宙機などの工学システムにおいて、最も重要な指標の一つである。安定性とは、システムに釣り合い状態からの逸脱が生じた際、一定時間経過後にシステムが元の釣り合い状態へ戻る性質を持っているかどうか、を定義した概念である。このような観点は、工学システムのみでなく、経済や生態活動など、自然科学や社会科学の様々な対象の解析においても、重要な役割を果たしている。

特に工学システムにおいては、システムの不安定は、過負荷状態の連鎖を引き起こして異常の原因になる事が考えられるため、安定性を監視する事は極めて重要である。安定性解析の起源は、システム制御工学の分野において 19 世紀後半頃にまで遡ることができるが、その議論の中心はシステムが安定であるための‘条件’に関するものが中心であり、与えられたシステムが安定であるか不安定であるかを判別する事に主眼が置かれている。しかし運用状態にあるシステムの監視においては、不安定な運用の結果である異常を捉えるのではなく、不安定状態へと向かう変調の前兆を捉える必要がある。従って筆者らは、システムに発生する変調の前兆の早期発見を行うためには、時系列データから‘安定性の変化’を捉える事が重要であると考える。そこで本稿では、まず安定性の観点から時系列データを比較する計量を定義する。一般に、安定性は状態空間モデルに基づく概念であるが、得られた計量は、システムの可観測行列の列ベクトルで張られる部分空間の距離として定義される。この計量を用いて、入出力系列から逐次的に学習される状態空間モデルの安定性の変化を評価する事ができるようになる。

本稿で述べる枠組みは、一つの変化点検知へのアプローチであるとも言える。変化点検知は、特にデータマイニングの分野で盛んに研究が行われており、様々なアプローチが提案されている。しかしその多くは、観測系列のみの時系列データを扱っているものが中心である [Markou 03a, Markou 03b, Ide 04, Fujimaki 06, Takeuchi 06]。本稿では、観測系列のみでなく、入力系列も含んだより一般な時系列データを対象とする、状態空間モデルに基づいた枠組みを提案している。実用場面では、工学システムに限らず、得られる時系列データが入力と出力に

分類可能である場合も多く、本稿で提案する枠組みは、このような対象へのより本質的なアプローチを可能とするという点においても、重要であると言える。

本稿の構成は、以下のようである。まず 2. では、本稿に必要な背景知識として、システムの安定性と部分空間計量の概要について説明する。そして 3. においては、可観測行列の部分空間計量が動的システムの安定性を比較する指標になっている事について述べる。更に 4. では、オンライン状態空間学習を用いて、部分空間計量に基づく安定性の比較を異常検知問題へと適用する方法について述べ、続く 5. において、具体的な検証実験によりその有用性を確かめる。最後に 6. では、本稿の結論と今後の展望について述べる。

2. 動的システムの安定性と部分空間計量

本章では、以降の章で必要となる背景知識として、動的システムの安定性と部分空間計量について、各々概要を説明する。

2.1. 動的システムの安定性

動的システムから得られる入出力系列 $\{u(t) \in \mathbb{R}^p, y(t) \in \mathbb{R}^q, t = 1, \dots, N\}$ を表現するモデルとして、次式で表される離散時間線形状態空間モデルを考える。

$$\begin{aligned} x(t+1) &= Ax(t) + Bu(t) + v(t) \\ y(t) &= Cx(t) + Du(t) + w(t) \end{aligned} \quad (1)$$

ここで、 $x \in \mathbb{R}^n$ は状態量であり、 $A \in \mathbb{R}^{n \times n}$ 、 $B \in \mathbb{R}^{n \times p}$ 、 $C \in \mathbb{R}^{q \times n}$ および $D \in \mathbb{R}^{q \times p}$ はシステムを表す定行列である。また、 $v \in \mathbb{R}^n$ および $w \in \mathbb{R}^q$ は各々白色雑音であり、システムノイズおよび観測ノイズと呼ばれる。本稿では以降、システム行列 A, B, C, D で決まる状態空間モデルを $M(A, B, C, D)$ と表す。このとき、システムの安定性の定義の一つ (正確には漸近安定性と呼ばれる) は、次のように与えられる。

定義 1 任意の初期状態 $x(0)$ に対して $\lim_{t \rightarrow \infty} \|x(t)\| = x_b$ 。ただし x_b は、システムの任意の平衡点である。

つまりシステムが安定であるとは、システムの状態が平衡状態から離れた場合に、時間の経過によって状態が発散しない事を

定義したものであり、定義から容易に分かるように、システム行列 A の固有値によって完全に決定される。ただしこの定義は、システムへの入力 u が全区間に渡って 0 である場合の挙動に関するものとなっている。従って本来は、更に入力に対してシステムがどのような挙動をしめすか、という点を考える必要があるように思われる。このような観点から定義された安定性は、入出力安定性と呼ばれ、伝達関数 $G(s)$ の極と呼ばれる量によって完全に決定される（一般に、入出力安定であれば漸近安定となる）。しかし後で述べるように、モデルがデータから推定されたものである場合、伝達関数 $G(s)$ の極と行列 A の固有値は完全に一致し、これらの安定性は等価となる。

2.2. 部分空間の正準角と距離

実行列 $A \in \mathbb{R}^{m \times p}$, $B \in \mathbb{R}^{m \times q}$ の列ベクトルで張られる部分空間を各々 S_1, S_2 とする。ただし、 A および B のランクを各々 r_a, r_b とした時、 $r_a < r_b$ であるとする。このとき、部分空間 S_1, S_2 の間の部分空間角は次のように定義される。

定義 2 (正準角 [Hotelling 36]) 部分空間 S_1, S_2 の間の部分空間角 $\theta_i \in [0, \pi/2]$ ($i = 1, \dots, r_a$)、および関連する正準方向 $u_i \in S_1, v_i \in S_2$ は、再帰的に次のように定義される。

$$\cos \theta_k = \max_{u \in S_1, v \in S_2} u^T v, \quad (2)$$

ただし $u^T u = v^T v = 1$ であり、かつ $k > 1$ において $u^T u_i = 0$ および $v^T v_i = 0$ ($i = 1, \dots, k-1$) を満足する。

つまり正準角のコサインは、行列の組 (A, B) の正準相関と等価なものである。従って、一般の正準相関解析と同様、特異値分解を用いて計算する事ができる。本稿では、正準角の組を

$$(\theta_1, \theta_2, \dots, \theta_{r_a}) = A \triangleleft B \quad (3)$$

と表す事にする。この正準角を用いて、部分空間の間の距離を次のように定義する事ができる。

定義 3 (部分空間計量 [Golub 89, Martin 00]) 部分空間 S_1, S_2 間の Gap d_g , Frobenius ノルム d_f , および Martin のノルム d は、各々次のように定義される。

$$d_g^2(S_1 - S_2) = \sin \theta_{max} \quad (4)$$

$$d_f^2(S_1 - S_2) = 2 \sum_{i=1}^{r_a} \sin^2 \theta \quad (5)$$

$$d^2(S_1 - S_2) = -\ln \prod_{i=1}^{r_a} \cos^2 \theta \quad (6)$$

文献 [De Cock 02] などから導かれるように、 d_f および d は、行列 A, B の QR 分解 $A = Q_1 R_1, B = Q_2 R_2$ で得られる行列 Q_1, Q_2 を用いて、次のように計算する事ができる。

$$d_f^2(S_1 - S_2) = -\ln \det(Q_1' Q_2)^2 \quad (7)$$

$$d^2(S_1 - S_2) = 2n - 2\text{tr}(Q_2' Q_1 Q_1' Q_2) \quad (8)$$

3. 部分空間計量による安定性の比較

本章では、入出力系列を生成している背後にある動的モデルを、安定性の観点から比較するための計量について考える。

一般に、利用可能なデータからなんらかの情報を抽出する際、「データから推定可能な限界はどこまでなのか？」という

点について考える事は、有用である場合が多い。この問いに対する答えは、いかなる問題に対しても与えられるものではないが、本稿で扱う問題では明確に答える事が可能である。つまり、入出力系列から推定される状態空間モデル (1) は、データを生成している真の状態空間モデルのうち、可観測かつ可制御な部分のみである事を示す事が可能であり*1、従って、「可観測かつ可制御という仮定の下で、状態空間モデルを比較する計量を導けば十分」となる。この仮定の下では、伝達関数 $G(s)$ の極は、システム行列 A の固有値と完全に一致するため、状態空間モデル (1) の安定性は、システム行列 A の固有値によって完全に決まる。今、システム行列を用いて、次式で表される行列 (可観測行列と呼ばれる) を考える。

$$\mathcal{O}^T = [C^T \quad (CA)^T \quad (CA^2)^T \quad \dots]^T \quad (9)$$

このとき、簡単な証明によって、可観測行列 \mathcal{O} は、システム行列 A の固有値 $\alpha_1, \alpha_2, \dots, \alpha_n$ を用いて表される行列

$$\Gamma = \begin{bmatrix} 1 & \dots & 1 \\ \alpha_1 & \dots & \alpha_n \\ \alpha_1^2 & \dots & \alpha_n^2 \\ \vdots & \ddots & \vdots \end{bmatrix} \quad (10)$$

と次のような関係を持つ事が示される*2。

$$\text{range}\{\mathcal{O}\} = \text{range}\{\Gamma\} \quad (12)$$

つまり、可観測行列の列ベクトルで張られる部分空間は、システム行列 A の固有値とその累乗から成るベクトルで張られる部分空間と一致する。従って、データから推定されたモデルを用いる場合、拡大可観測行列 (9) は、システムの安定性に関する全ての情報を持っている事がわかる。そこで本稿では、安定性の観点から状態空間モデル (1) を比較する計量として、拡大可観測行列に関する部分空間計量を用いる。

定義 4 入出力系列から推定された 2 つの状態空間モデルの組を $SSM_1(A_1, B_1, C_1, D_1), SSM_2(A_2, B_2, C_2, D_2)$, また各々のシステム行列を用いて定義される拡大可観測行列を $\mathcal{O}^{(1)}, \mathcal{O}^{(2)}$ とする。このとき、モデル SSM_1, SSM_2 間の部分空間角および部分空間計量を次のように定義する。

$$SSM_1 \triangleleft SSM_2 := \mathcal{O}^{(1)} \triangleleft \mathcal{O}^{(2)} \quad (13)$$

$$d_g(SSM_1, SSM_2) := d_g(\mathcal{O}^{(1)} - \mathcal{O}^{(2)}) \quad (14)$$

$$d_f(SSM_1, SSM_2) := d_f(\mathcal{O}^{(1)} - \mathcal{O}^{(2)}) \quad (15)$$

$$d(SSM_1, SSM_2) := d(\mathcal{O}^{(1)} - \mathcal{O}^{(2)}) \quad (16)$$

安定性は、前述のように状態量の軌跡に関する概念である。従って、可観測行列の計量が安定性の観点から動的システムを比較するものとなっている事は、任意の初期状態 x を与えたときの、その後の状態量の軌跡が可観測行列によって一意に決められるという事実からも確認する事ができる。

*1 詳細に関する説明は制御理論の教科書に譲る。

*2 行列 A はフルランクなので、正則行列 T により $A = T\Lambda T^{-1}$ ($\Lambda = \text{diag}(\alpha_1, \dots, \alpha_n)$) と対角化可能である。このとき、

$$\mathcal{O}^T = (\text{diag}(CT) \cdot \Gamma \cdot T^{-1})^T \quad (11)$$

となり、 $\text{diag}(CT)$ および T^{-1} がフルランクである事から、 \mathcal{O} の列ベクトルが張る部分空間と、 Γ のそれは一致する。

4. 異常検知問題への適用

異常検知問題において扱うべき異常は、大きく (1) 突発的变化、および (2) 漸化的変化の二種類の変化から生じるものに分類できる。工学システムであれば、例えば、前者はなんらかの衝撃等により瞬間的な構造的変化や機能的変化が生じた場合に相当し、また後者は機器の磨耗等によるパラメータ変化が生じる場合に当たる。容易に想像がつく事であるが、一般に、異常検知の立場からすると、前者の方が問題としては扱いやすい。しかし冒頭でも述べたように、重要な事は、システム変調の前兆を早期に発見する事であるので、異常検知の枠組みとしては、後者の変化を適切に捉えられる事が重要となる。

異常検知へのアプローチとしては、次の二つが考えられる。

1. 参照区間と現時点でのモデルとの比較。
2. 一定の時間差を持つ区間間でのモデルの比較。

1 のアプローチは、多くの従来研究で採用されてきたアプローチで、通常は正常と思われる区間を参照区間として選択し、その区間のデータを用いて得られた、いわゆる‘正常モデル’と、現時点で得られているモデルとを比較する事により、有意な変化を検知するものである。また 2 のアプローチは、システムが時変、つまり時間と共にデータの背後にあるシステムが変化していると仮定できるような対象に対しては、特に有用である [Takeuchi 06]。しかし穏やかな変化を伴う異常に対しては、推定されたモデルが異常を含めんだモデルとなってしまう、適切に検知できなくなってしまうという問題がある。つまり、いずれのアプローチも長所短所があるため、個々の問題や状況に応じて、どちらのアプローチを選択するかを適切に判断する必要がある。従って重要なのは、異常診断の枠組みが、いずれのアプローチにも対応できる事であると言える。

本稿ではここまで、安定性の観点から入出力系列の背後にある動的システムを比較する計量について述べてきた。この計量は、式 (1) で表される状態空間モデルに基づくものである。従って、時系的に得られる入出力系列を用いて逐次的に状態空間モデルを学習する事により [Lovera 00]、得られたモデルと、ある参照区間のデータを用いてオフラインで学習されたモデル、あるいは一定時刻前に推定されたモデルとの距離を、この計量を用いて評価する事が可能となる。異常検知への適用においては、計量の評価のみが必要となるため、データからモデルを推定し、そして得られたモデルから計量を計算する、という 2 段階のステップは非効率であるとも言える。しかし現時点では、逐次的に上記 2 つのアプローチを可能とするには、これらの手順が必要であると考えている。なお、後述の検証実験においても、両アプローチによる結果を示す。

5. 実験と考察

異常検知問題における検証のため、静止衛星の軌道上における姿勢マヌーバ・シミュレーションのデータに対して、前章で述べたアプローチを適用した結果を示す。

本シミュレーションの概要は、次のようである。まず模擬される衛星は、リアクションホイール (RW) と呼ばれる制御機器を 3 機用いてフィードバック制御により一定の姿勢に保たれており、得られるデータとしては、姿勢角 (3) と姿勢角速度 (3)、および RW の回転数 (3) の観測値計 9 系列、フィードバック制御値として RW への姿勢制御トルク 3 系列、またプラントおよび制御器への白色同定信号が各々 6 系列および 3 系列である。またシミュレーションのストーリーとしては、まず初期時刻から 1,500[s] まで一定の姿勢 (全ての姿勢角および角

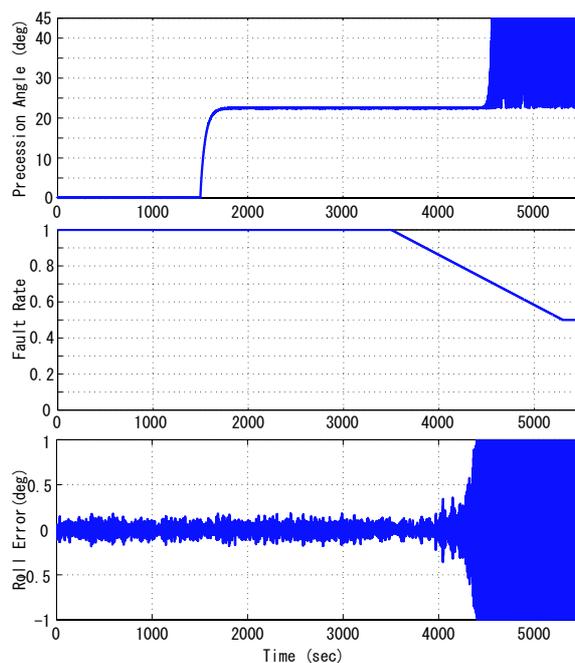


図 1: 実験に用いたシミュレーション・データの一部

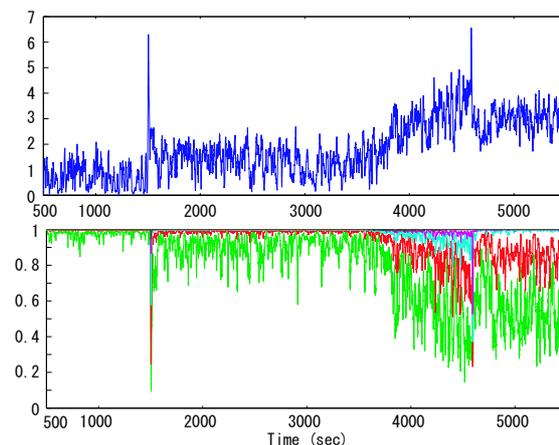


図 2: 実験結果 (時間差を持つ 2 モデル間の部分空間計量 (Martin ノルム:上) と部分空間角のコサイン:下)

速度が 0) に制御されていた衛星に対し、時刻 1,500[s] に姿勢変更 (歳差角 22.5 への変更。図 1 の第 1 図を参照) が実行される。そして姿勢変更の実行後は、その姿勢を維持する姿勢制御を続けるが、時刻 3,500[s] から RW のうちの一つ (Roll 軸を制御する RW) に異常が発生するというものである。発生する異常は、図 1 の第 2 図に示すように、時刻 3,500[s] から 1,300 秒間かけて、規定値の半分の出力しか出ない状態まで漸化的にパラメータが変化するというものである。またデータの取得間隔は 0.25[s] であり、全体で 5500 秒間 (従ってデータは 22,000 点) のシミュレーションとなっている。なお、図 1 の第 3 図は Roll 軸まわりの姿勢角の制御誤差を図示したものであり、この図および同第 1 図を見ても分かるように、前述の異常によるシステムへの影響は、異常発生から約 5,00 秒後の時刻 4,000[s] 付近から現われ始め、時刻 4,500[s] 直前頃から、衛星は完全に姿勢制御機能を損失している事が分かる。

本シミュレーション・データに対して、最初の 5,00 秒間 (2,000 データ点) を初期モデルの推定に用い、それ以後をオンライン

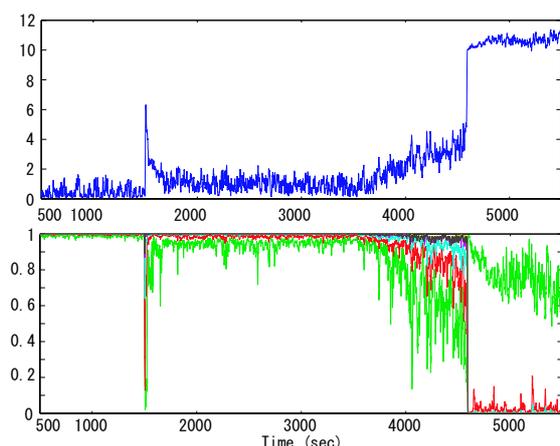


図 3: 実験結果 (参照モデルと現時点モデル間の部分空間計量 (Martin ノルム:上) と部分空間角のコサイン:下)

学習と 3. で述べた動的システムの計量により, その安定性の変化を評価したものが図 2 および図 3 である. 図 2 は 5 秒間 (20 データ点) の時間差を持った 2 モデル間の部分空間角または距離 (Martin ノルム) であり, また 3 は初期推定で得られたモデルと逐次的に推定されたモデル間の部分空間角または距離である. いずれの場合も, 用いた状態空間の次元は $n = 7$ (従って Martin ノルムの最大値は 14), ハンケル行列のブロック数は 25, また忘却率は 0.98 である. いずれの結果も, 時刻 1,500[s] における姿勢変更開始時 (突発的变化), および時刻 3,500[s] からの漸化的異常に対して, 各々の計量が瞬間的あるいは漸的に増加しており, 適切に入出力データの背後にある動的モデルの変化を捉えられている事がわかる. 特に漸化的異常に関しては, パラメータの変化がシステムの挙動を大きく変えるものではなく, 前述のように観測値などに影響が出るのはかなりの時間遅れを伴うにもかかわらず, 変化開始の早期からその変化を検出できている. これは, 本稿で導いた計量がデータの背後にあるモデルの性質から導かれたものであるため, データの本質的な情報を適切に抽出できている結果であると思われる.

6. 結論

本稿では, 逐次的に入出力系列から推定される状態空間モデルの可観測行列に関する部分空間計量が, データの背後にある動的システムの安定性を比較する計量となっている事を示した. そして, この計量を異常検知問題へ適用する際のアプローチについて述べると共に, その検証のために, 静止衛星の軌道上における姿勢マヌーバ・シミュレーションのデータに対して適用し, その有用性を示した. また本稿で述べた枠組みは, 入出力系列データを対象とする状態空間モデルに基づいたものであり, 得られる時系列データが入力と出力に分類可能である対象へのより本質的な変化点検知へのアプローチを可能とするという点においても, 重要な提案であると言える.

今後の課題として, 筆者らは次のような問題を考えている. 本稿では, 状態空間モデルを比較する計量を導き, その変化点検知における有用性を示した. しかし計量を評価するためには, まずデータからモデルを推定し, 得られたモデルから計量を計算する, という 2 段階のステップが必要となる. 変化点検知などの応用などにおいて計量の計算のみが必要となる場合, この枠組みは非効率であると言える. 従って筆者らは今後, 今回得られた洞察を元に, データから直接評価が可能な計量の導

出, およびその逐次計算の方法について, 検討を進めたいと考えている.

参考文献

- [Markou 03a] Markou, M. and Singh, S.: Novelty Detection: A Review - Part 1: Statistical Approaches, *Signal Processing* **83**(12): 2481-2497 (2003).
- [Markou 03b] Markou, M. and Singh, S.: Novelty Detection: A Review - Part 2: Neural Network Based Approaches, *Signal Processing* **83**(12): 2499-2521 (2003).
- [Takeuchi 06] Takeuchi, J. and Yamanishi, K.: A Unifying Framework for Detecting Outliers and Change Points from Non-Stationary Time Series Data, *IEEE Transactions on Knowledge and Data Engineering* **18** (4): 482-489 (2006).
- [Ide 04] Ide, T. and Kashima, H.: Eigenspace-based Anomaly Detection in Computer Systems, *Proc. of the 10th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* (2004).
- [Fujimaki 06] Fujimaki, R., Yairi, T. and Machida, K.: An Approach to Spacecraft Anomaly Detection Problem Using Kernel Feature Space, *Proc. of the 12th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* (2006).
- [Hotelling 36] Hotelling, H.: Relations Between Two Sets of Variates, *Biometrika* **28**: 321-372 (1936).
- [Golub 89] Golub, G. H. and Van Loan, C. F.: Matrix Computations (2nd eds.), *Johns Hopkins University Press* (1989).
- [Martin 00] Martin, R. J.: A Metric for ARMA Processes, *IEEE Trans. on Signal Processing* **76**:93-97 (2000).
- [De Cock 02] De Cock, K. and De Moor, B.: Subspace Angle between ARMA Models, *System & Control Letters* **46**: 256-270 (2002).
- [Lovera 00] Lovera, M., Gustafsson, T. and Verhaegen, M.: Recursive Subspace Identification of Linear and Non-linear Wiener State-Space Models, *Automatica* **36**: 1639-1650 (2000).