

ダイナミクス法則式及び隠れ状態の発見とカオス同定への適用

Discovery of Dynamic Law Equations and Hidden States, and Identification of Chaotic Process

足立 史宜*¹ 鷲尾 隆*¹ 元田 浩*¹
 Fuminori ADACHI Takashi WASHIO Hiroshi MOTODA

*¹大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

This paper proposes a novel approach named SCALETRACK to discover dynamic models represented by simultaneous time differential law equations including hidden states from time series data measured in an objective process. The power of SCALETRACK comes from correlation dimension analysis to detect hidden states, scale-type constraints to limit the search to equations representing first principles, and nonlinear state tracking to evaluate the fitness of candidate equation formulae. The promising performance of the proposed approach is demonstrated through the analysis of synthetic data on engineering process and chaotic behaviors.

1. はじめに

これまで、科学者たちは実験や観測によって得られた数値データに内包されている規則性を読み取り、“科学的法則式”を発見し、それらを組み合わせることによって複雑な現象の“科学的モデル式”を構築してきた。従来よりこの作業に用いられてきた手法は、システム同定理論 [Ljung 86] や、ニューラルネットワーク [Wasserman 89] である。これらの手法は狭い状態の範囲内で漸近的に成り立つモデルを構築し、そのモデルが広い範囲で一般的に成立する事を仮定している。しかし、構築されたモデルは狭い状態の範囲内で仮定されたものであり、広い範囲にわたって一般的に成立するモデルであることは少ない。したがって、これらの手法では、一般性を有する第1原理に基づいた科学的法則式を発見することは困難である。

これに対して、科学的法則式を発見を目的として、Langley等による BACON [Langley 87] が提案された。BACON は対象の状態を任意の値に制御可能であるという“能動的観測”の条件で取得されたデータの中から、その中に内包される完全方程式*¹を探索する。BACON は、状態変数のうち、任意の2数量組に対して実験により得られるデータへのあてはめを行い、2数量組の関係を多様な候補の中から選択する。さらに多数の2数量組の関係同定や中間変数の作成を繰り返しボトムアップに多変量からなる完全方程式を組み上げる。同様の手法を用いたものに FAHRENHEIT [Koehn 86], ABACUS [Falkenhainer 86], IDS [Nordhausen 90] などのシステムが存在し、これらは BACON ファミリーと呼ばれている。これらのシステムで発見された関係式は、一般に与えられたデータの範囲内において成立するが、発見された関係式が数学的に許容される式であるという保証はない。これに対して鷲尾等は、スケールタイプ制約を始めとするいくつかの制約を用いて、関係式の候補を第1原理に基づいたものに限定した SDS [Washio 97a] を提案した。しかし、実際の電気回路のような対象は、対象の中にある複数のメカニズムを表す連立方程式によってその振る舞いが記述され、1つの完全方程式では記述しきれない。この問題に対し、連立方程式の発見手法として、SSF [Washio 98] が提案されている。SSF は実験により測定された状態量の最小完全集合を

連絡先: 足立 史宜, 大阪大学産業科学研究所, 〒567-0047
 大阪府茨木市美穂ヶ丘 8-1, TEL/FAX 06-6879-8544,
 adachi@ar.sanken.osaka-u.ac.jp

*¹ n 個の変数からなり、 n 個の変数の値のとり方に $n-1$ の自由度がある方程式。

同定することで対象としているプロセスを記述する連立方程式を発見する。

しかし、上に挙げた手法は、静的かつ単純な電気回路のように、能動的観測が可能な条件下でしか法則式を発見することはできない。つまり、対象のどの状態も変更不可能な“受動的観測”において採取されたデータから法則式を発見することはできない。この問題に対し、拡張 SDS や拡張 SSF はデータからサンプリングを行って擬似実験を行うことで受動的観測データからの法則式発見を可能にしている [Washio 99, Washio 01]。さらに、受動的観測時系列データからの連立微分方程式を発見する試みが Dzeroski らによりなされ、LAGRANGE [Dzeroski 94] が提案されている。LAGRANGE は受動的に観測された時系列データから、ILP に似た仮説の生成と試験を行うことにより連立微分方程式の発見を行う。また、LAGRANGE の拡張である LAGRAMGE [Todorovski 97] は対象に関する領域知識を導入し、候補式の探索空間を減らすと共に、発見された式の信頼度を高めている。

しかし、科学者は、領域知識が使える対象なら、法則式発見システムに頼ることなく良い式を構築することが可能である場合が多い。したがって、法則式発見システムの需要は、領域知識が不十分な状態において人間が構築困難な式を発見するところにある。そのような例としては、“隠れ状態”を含む連立微分方程式の発見がある。LAGRANGE や LAGRAMGE は受動的観測の条件下で連立微分方程式を発見することができるが、すべての状態量が観測可能であることを前提としており、隠れ状態がある対象についての法則式を得ることができない。

本論文では、SCALETRACK (SCALE-type and state TRACKing based discovery system) と名づけた以下の特徴を持つ法則式発見手法を提案する。

- (1) 発見する方程式は対象の動的な振る舞いを反映する連立微分方程式である。
- (2) 発見する方程式は第1原理にを表す可能性の高い法則式である。
- (3) 対象内に存在する隠れ状態を特定し、その振る舞いを記述する方程式も発見する。
- (4) 対象を限定する領域知識は使わない。
- (5) 受動的観測データからの法則式発見を行う。

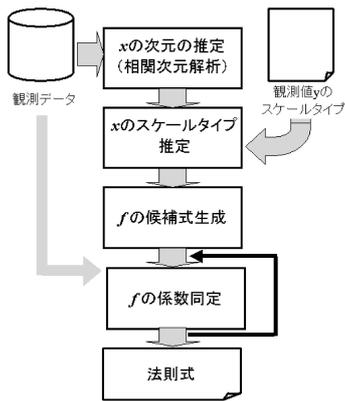


図 1: SCALETRACK のブロック図

本論文の第 2 節では問題設定と SCALETRACK の概要について述べ、また、第 3 節で性能検証のための実験と結果、及びカオスへの適用例について述べる。

2. SCALETRACK の概要

2.1 問題設定

SCALETRACK は、状態変化と観測に関して以下の“状態空間モデル”を採用する。

$$\begin{cases} \dot{x}(t) = f(x(t)) + v(t) \\ y(t) = Cx(t) + w(t) \end{cases}$$

$(v(t) \sim N(0, \Sigma_v), w(t) \sim N(0, \Sigma_w))$

ここで、第 1 番目の式を“状態方程式”、第 2 番目の式を“観測方程式”と呼ぶ。また、 x は“状態変数ベクトル”、 $f(x)$ は“状態関数”、 v は“状態ノイズベクトル”、 w は“観測ノイズベクトル”、 y は“観測変数ベクトル”、 C は“観測行列”、 t は離散時間を表すインデックスである。 $f(x)$ は一般的に非線形の関数であり、 x についての任意の状態遷移を表すことができる。また、 C は観測装置を表す行列で、状態変数ベクトルから観測変数ベクトルへの変換を表すものである。観測装置は人為的に作成され、通常線形であるので、この状態空間モデルは一般性を損なわない。もし、 C が単位行列であった場合、すべての状態変数が直接観測可能であることを表す。また、 C が列フルランクであった場合、観測方程式を x について解くことによって観測誤差の範囲内で状態変数の値を求めることができる。そのほかの場合には、観測方程式からだけでは、値を特定できない状態変数が存在する。そのような変数が前述した隠れ状態を表す“隠れ状態変数”である。

実際の発見プロセスにおいては、 $f(x)$ と、 x のいくつかの要素は未知であり、観測装置により観測される部分ベクトルである $x' (\subseteq x)$ の存在と部分観測行列 $C' (\subseteq C)$ だけしか知ることができない。つまり、SCALETRACK は、初めに隠れ状態を含めた状態変数ベクトルの次元を求める必要がある。そして、その後第 1 原理を反映した $f(x)$ を探索する。

2.2 概要

SCALETRACK の処理ブロック図を図 1 に示す。システムに与えられるのは観測値の時系列データと観測変数のスケールタイプに関する知識である。SCALETRACK は、まず相関次元解析 [Berge 92] と呼ばれる統計解析手法を用いて観測デ

ータから状態ベクトルの次元を推定する。相関次元解析は、観測ベクトルのある要素とその時間遅れ要素から構成される相空間上に、観測ベクトルを射影し、その軌跡の自由度を計算することによって状態変数ベクトルの次元を推定するものである。

状態変数ベクトルの次元が推定されると、与えられた観測変数のスケールタイプと部分観測行列 C' から、スケールタイプに基づいて、状態変数のスケールタイプについてすべての可能な組み合わせを列挙する。ここで、スケールタイプは観測の性質により分類されるもので、物理量に関しては主に“比例尺度”と“間隔尺度”の 2 つに分類される。比例尺度量は絶対的な原点を持つ。また、許容される単位変換が相似変換 ($x' = cx$) であるので、2 つの測定量の比は単位にかかわらず一定である。比例尺度量の例として、質量や速度、周波数などがある。一方、間隔尺度量は任意に設定された原点を持つ。また、許容される単位変換は一般線形変換 ($x' = cx + z$) である。間隔尺度量の例としては、摂氏温度、エネルギー、エントロピー、座標等が挙げられる。比例尺度量と間隔尺度量に関して、それぞれの量が直接の依存関係を持つ場合には、その関係が各数量のスケールタイプにより決まる基礎的な関数で表される。この関係を記述したものがスケールタイプ制約である。

続いて、得られたすべての状態変数のスケールタイプの組み合わせについて、数学的に許容される $f(x)$ の候補式を作成する。このときに用いるのは、“Extended Product Theorem [Washio 97b]”である。Extended Product Theorem は、2 数量間の関係の制約を表したスケールタイプ制約を多数量間の関係を表すように拡張したものである。

さらに、“Sequential Importance Sampling/Resampling Monte Carlo filter (SIS/RMC filter) [Doucet 00]”を用いて、与えられた観測データと候補式を当てはめて状態変化をシミュレーションし、候補式の係数を最適化する。SIS/RMC フィルタは、与えられた状態方程式と観測値から、状態変数の確率分布を作成し、サンプリングを行うことにより N 個の系列と重み変数の組を作成し、確率の高い (つまり重み変数の値が大きい) 状態変数の系列を推定するフィルタである。観測データと状態変化のシミュレーションにより得られた疑似観測値の間の平均二乗誤差の意味で、もっとも精度の高い候補式と係数の組み合わせが出力される。このように SCALETRACK は受動的観測データから、スケールタイプに関する知識と観測装置以外の領域知識を使わずに、第 1 原理に基づいた状態方程式を発見する。

3. 実験

3.1 実装

SIS/RMC フィルタによる推定は膨大な計算時間を必要とする処理である。しかし、探索空間は非連続、非単調であるためどの処理も省略することはできない。単体の計算機で SCALETRACK を実行した場合、解を得るまでに 1ヶ月以上の時間が必要である場合もある。そこで、SCALETRACK は 1 台のサーバーと 10 台のクライアントを結合したクラスター計算機を用いて実装した。サーバーは CPU に AthlonXP1900+ を搭載し、メインメモリは 2GB である。また、クライアントは CPU に AthlonXP3000+ を搭載し、メインメモリは 512MB である。サーバーは図 1 に示された最初の 3 ステップ (相関次元解析、 x のスケールタイプ推定、候補式生成) を行い、候補式と 1 つのべき級数候補組の組み合わせを 1 つのワークユニットとして、クライアントに分配する。クライアントは比例係数組の最適化を行い、結果をサーバーに返す。ワークユニットは

表 1: 各ノイズレベルにおける $\nu_{max}(8)$ の値

case	$\sigma_w(\%)$					
	0.1	0.5	1.0	2.0	5.0	10.0
RR	1.11	1.74	2.21	2.59	3.23	3.44
RRH	1.05	1.74	2.21	2.59	3.23	3.44
RI	1.25	1.90	2.19	2.40	2.41	2.58
RIH	1.25	1.90	2.19	2.40	2.41	2.58

表 2: 基本性能評価

case	ct (hrs.)	$\sigma_w(\%)$				
		0.1	0.5	1.0	2.0	5.0~
RR	1.5	++	+	+	+	-
RRH	5.5	+	+	-	-	-
RI	4.0	++	+	+	+	-
RIH	5.5	++	+	-	-	-

独立で、他のワークユニットと干渉しないため、この結果、処理速度はほぼ 10 倍になる。

3.2 基本性能

以下の 2 つのモデルについて人工データを作成し、SCALETRACK の基本性能を検証した。

1. Model RR:

$$\begin{cases} \dot{x}_1(t) &= x_1(t)x_2(t) \\ \dot{x}_2(t) &= -0.5x_2(t) \end{cases}$$

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + w_t$$

ただし、 $y_1(t) = x_1(t)$ and $y_2(t) = x_2(t)$ は共に比例尺度変数である。また、観測値のタイムステップの幅は $\Delta_t = 0.005$ で、データの総ステップ数は 600 である。

2. Model RI:

$$\begin{cases} \dot{x}_1(t) &= 0.4x_1(t)(x_2(t) + 0.2) \\ \dot{x}_2(t) &= -0.1(x_2(t) + 0.6) \end{cases}$$

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + w_t$$

ただし、 $y_1(t) = x_1(t)$ は比例尺度変数で、 $y_2(t) = x_2(t)$ は間隔尺度変数である。また、観測値のタイムステップの幅は $\Delta_t = 0.05$ で、データの総ステップ数は 600 である。

各モデルの観測ノイズ w_t は、 $w_t, x(t)$ の h 番目の要素をそれぞれ $w_t^h, x_h(t)$ とすると以下のように表される乱数である。

$$w_t^h \sim N(0, \sigma_w x_h(t))$$

ただし、 σ_w は状態変数の値に対するノイズの強さである。今回は、相関次元解析の相空間の次元 m を 8、SIS/RMC フィルタの総数 N を 500 とした。また、純粋な観測ノイズの影響を測定するために状態ノイズベクトルの値は全て 0 にしている。

表 1 は相関次元解析の結果を示している。RR, RI はそれぞれ上記のモデル RR と RI のデータに対応している。また、RRH, RIH は RR, RI それぞれのモデルから、2 番目の状態変数を隠れ状態としたものである。この結果より、相関次元解析の値は多くの場合でほぼ 2 付近となっており、2 次元の状態空間を仮定して探索を行った。表 2 は性能評価の結果を表している。ct の列は計算時間を表している。RRH, RI, RIH の計算に必要な時間は RR の計算にかかる時間よりも長い。これは、間隔尺度変数を含む候補式は比例尺度変数のみの候補式と比べて式の数が多く、同定する係数の数も多くなるからである。表中で ++ と表記されているものは、最も確度が高いと評価された式が正しい式であったもので、+ と表記しているものは、正しい式が上位 5 つの式に含まれていたものである。-

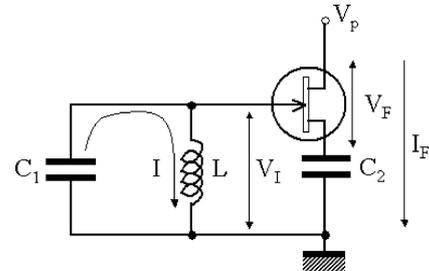


図 2: 使用した電気回路図

方、- と表記されているものは、上位 5 つの中にも正しい式が含まれなかったものである。表 2 によると、隠れ状態が無いモデルはノイズの強度が 2.0% までは正しい式を出力していることが分かる。また、隠れ状態がある場合にも、ノイズの強度が 0.5% までは正しい式を発見している。

3.3 電気回路のダイナミクスの発見

基本性能の評価に加えて、電気回路をシミュレートしたデータに対して SCALETRACK を適用した。電気回路は図 2 に示されるような LC と電界効果トランジスタ (FET) からなる単純な回路である。この回路のダイナミクスは以下の式で表される。

$$\begin{cases} \dot{V}_I(t) &= -\frac{I(t)}{C_1} &= -100I(t) \\ \dot{I}(t) &= \frac{V_I(t)}{L} &= 50V_I(t) \\ \dot{V}_F(t) &= \frac{V_I(t)V_F(t)}{rC_2} &= 250.0V_I(t)V_F(t) \end{cases}$$

ただし、式中の $V_I, I, V_F, L = 20\text{mH}, C_1 = 10\text{mF}, C_2 = 1\text{mF}$ は図 2 の記号と対応し、 $r = 4.0\Omega V$ は FET の電圧抵抗係数である。このとき、状態変数のスケールタイプは全て比例尺度である。観測値の時系列データは時間ステップ幅 $\Delta t = 0.001$ で作成され、総ステップ数は 800 である。また観測ノイズの強度は 0.1% としている。相関次元解析の結果、 $\nu_{max}(8) = 2.94$ が得られたので、状態変数の個数を 3 個と仮定して式の探索を行った。

まず、全ての変数が直接観測可能、つまり、観測行列 C が 3 次元の単位行列であるときは、最良の解として、以下の方程式が出力された。

$$\begin{cases} \dot{V}_I(t) &= -133.3I(t) \\ \dot{I}(t) &= 6.94V_I(t)V_F(t) \\ \dot{V}_F(t) &= 249.0V_I(t)V_F(t) \end{cases}$$

この式は第 1 番目と第 3 番目の方程式の形が正しい式と同じであり、係数の値も比較的近い。

続いて、 I を隠れ状態として観測値に含めなかった場合には

表 3: カオス式の探索結果 (Z のみ)

順位	式	MSE
1	$\frac{dZ}{dt} = 0.001(X - 4.5)(X + 1000Z - 4.5)$	5.46×10^{-7}
2	$\frac{dZ}{dt} = 5.00(X + 0.627Z - 4.06)$	5.81×10^{-5}
3	$\frac{dZ}{dt} = 2.61(X + 0.32Y - 0.62)$	6.16×10^{-5}
4	$\frac{dZ}{dt} = 0.87(X - 0.19)$	6.74×10^{-5}
5	$\frac{dZ}{dt} = 1.06(X + 0.29Y + 1.12Z + 1.11)$	6.62×10^{-5}

上位の 5 つの式の中に以下の式が含まれていた .

$$\begin{cases} \dot{V}_I(t) &= -26.9I(t) \\ \dot{I}(t) &= 298.0V_I(t) \\ \dot{V}_F(t) &= 250.0V_I(t)V_F(t) \end{cases}$$

この式は、係数の値は第 3 式を除いて大幅に違った値となっているが、全ての式の形が正しい形と同じである . これらの結果は、SCALETRACK が少なくとも 3 次元の対象に対して適用可能であることを示している .

3.4 カオスデータに対する適用

さらに、以下のような変形レスラーカオスを表す状態方程式から得られるデータに対してスケールタイプによる候補式の作成と SIS/RMC フィルタを用いたトラッキングを行った .

$$\begin{cases} \frac{dX}{dt} = -Y - Z \\ \frac{dY}{dt} = X + 0.36Y \\ \frac{dZ}{dt} = 0.001(X - 4.5)(X + 1000Z - 4.5) \end{cases}$$

観測行列は 3 次元の単位行列であるとし、観測値の時系列データは時間ステップ幅 $\Delta t = 0.0001$ で作成し、総ステップ数は 1500 である . また、観測ノイズの強度は 0.1% としている . X , Y , Z の観測値は共に間隔尺度変数であるが、候補式の数が多いにも多く、全てを探索することは困難である . そこで、状態方程式の X と Y についての方程式を正しい式に固定し、 Z の式をスケールタイプ制約から得られる式のいずれかを当てはめた . Z に与えた式を、SIS/RMC フィルタで推定された Z の状態と観測値の時系列データの間の MSE の値が小さい順に表 3 に示す . 表 3 より、正しい式を与えたときの MSE は他の式の MSE 値より小さくなっており、スケールタイプ制約と Extended Product Theorem による候補式の生成と、SIS/RMC フィルタによる推定を組み合わせる方法を用いることで、カオスを表す方程式の同定を行うことが可能であることがわかる .

4. 結論

本論文では、受動的観測データから第 1 原理を反映する法則式を発見する手法について提案し、評価実験を行った . この手法は発見プロセスにおいて、スケールタイプ以外の対象に特化した領域知識を必要とすることなく、第 1 原理を反映する法則式を得る事が可能で、広い範囲の対象に対して適用することが可能である . また、従来の手法では扱えなかった隠れ状態を扱うことができ、さらに適用範囲の拡大が期待できる . 適用範囲を拡大するための計算量の問題の解決とノイズ耐性の強化、及びより広範囲に探索を行うための探索方法の改良は今後の課題である .

参考文献

[Berge 92] Berge, P., Vidal, C. and Pomeau, Y. (Aisawa, A., tr.): Order within Chaos - Towards a Deterministic Approach to

Turbulence, Sangyo-Tosho, ISBN: 4782800681, pp.133-147, 1992 (in Japanese).

- [Doucet 00] Doucet, A., Godsill, S. and Andrieu, C.: On Sequential Monte Carlo Sampling Methods for Bayesian Filtering, Statistics and Computing, Vol.10, Issue 3, pp.197-208, 2000.
- [Dzeroski 94] Dzeroski, S. and Todorovski, L.: Discovering Dynamics: From Inductive Logic Programming to Machine Discovery, Journal of Intelligent Information Systems, Kluwer Academic Publishers, pp. 1-20, 1994.
- [Falkenhainer 86] Falkenhainer, B. C. and Michalski R. S.: Integrating Quantitative and Qualitative Discovery: The ABACUS System, Machine Learning, pp.367-401, Boston, Kluwer Academic Publishers, 1986.
- [Koehn 86] Koehn, B. and Zytkow, J. M.: Experimenting and theorizing in theory formation, Proceedings of the International Symposium on Methodologies for Intelligent Systems, pp. 296-307, ACM SIGART Press, 1986.
- [Kokar 86] Kokar, M. M.: Determining Arguments of Invariant Functional Descriptions, Machine Learning, Kluwer Academic Publishers, pp.403-422, 1986.
- [Langley 87] Langley, P.W. et al.: Scientific discovery; Computational explorations of the creative process, Cambridge, MA: MIT Press, 1987.
- [Ljung 86] Ljung, L.: System Identification, P T R Prentice-Hall, 1986.
- [Nordhausen 90] Nordhausen, B. and Langlay, P.W.: An Integrated Approach to Empirical Discovery, Computational Models of Scientific Discovery and Theory Formation, Morgan Kaufman Publishers, Inc, San Mateo, California, 1990.
- [Takens 81] Takens, F.: Detecting strange attractors in turbulence., Lecture Notes in Mathematics, volume 898, pp. 366-381, Dynamical Systems and Turbulence, Warwick, 1981.
- [Todorovski 97] Todorovski, L. and Dzeroski, S.: Declarative Bias in Equation Discovery, Proceeding of the fourteenth International Conference on Machine Learning, Morgan Kaufmann, pp. 376-384, 1997.
- [Wasserman 89] Wasserman, P. D.: Neural Computing: Theory and Practice, New York: Van Nostrand Reinhold, 1989.
- [Washio 97a] Washio, T. and Motoda, H.: Discovering Admissible Models of Complex Systems Based on Scale-types and Identity Constraints, Proceedings of IJCAI'97, Vol.2, Nagoya, pp. 810-817, 1997.
- [Washio 97b] Washio, T. and Motoda, H.: Discovery of First Principle Based on Data-Driven Reasoning, Proceedings of the Fiest Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp.169-182, 1997.
- [Washio 98] Washio, T. and Motoda, H.: Discovering Admissible Simultaneous Equations of Large Scale Systems, Proceedings of AAAI'98, pp.189-196, 1998.
- [Washio 99] Washio, T., Motoda, H. and Niwa, Y.: Discovering admissible model equations from observed data based on scale-types and identity constraints, Proceedings of IJCAI'99, Vol.2, pp.772-779, 1999.
- [Washio 01] Washio, T., Motoda, H. and Niwa, Y.: Discovering Admissible Simultaneous Equation Models from Observed Data, Proc. Of 12th European Conference on Machine Learning:ECML2001, pp.539-551, 2001.