3F3-02

パ-ティクルフィルタを用いた法則微分方程式の発見

Discovery of Scientific Law Differential Equations with Particle Filters

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
I.S.I.R. Osaka University

In this paper, we propose an approach to discover scientific law differential equations which represents a dynamical system based on its observed time series data. This approach takes advantage of scaletype constraints to decide the style of admissible equations and particle filter to reduce influence of process noise and observation noise.

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．近年，計算機の高性能化や情報処理技術の発達と共に，人工知能や知識発見など広い分野に於いてこれらの行為を支援，自動化する試みが行われており，P.W.Langley 等による BACON[1][2] や，その派生である FAHRENHEIT[3]，ABACUS[4] などの実験的関係を発見するシステムが提案されている．これらのシステムで発見された実験的関係式は，一般に実験の範囲内においての客観性，再現性，普遍性，健全性を有するが，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．そこで，鷲尾等は BACON の考え方にスケールタイプ制約 (*Scale-type Constraint*)[5] をはじめとするいくつかの制約を加え，極力高い客観性，再現性，普遍性，健全性，数学的許容性を持つ法則式の発見が可能な “*Smart Discovery System*(SDS)” [6] を提案した．

スケールタイプとは，簡単に言えば単位が持つ性質を分類したもので，物理量に関しては２つの測定量の差を表す “間隔尺度”，絶対的な原点によって定められる “比例尺度”，いわゆる無次元量である “絶対尺度” の３つがある [7]．これらのうち，間隔尺度量および比例尺度量に関して，それぞれの量の間に絶対尺度量を媒介としない直接的な依存関係を持つ場合には，その関係は各数量のスケールタイプの性質によって決まる基礎的な関数で表される．これがスケールタイプ制約である [5]．

ここで，BACON 系のシステムや SDS は対象の状態を能動的に変更し観測することが可能であることが前提となっており，受動的観測しか許されない場合には適用することはできない．受動的に観測されるデータからの法則発見手法としては Dzeroski 等による LAGLANGE[8] 等が存在する．LAGLANGE は受動的かつ動的な時系列データを対象とした実験的関係の発見を行なうことが出来るが，BACON 系と同様に得られた実験的関係式が数学的許容性を満たしている保証はない．これに対して受動的観測データから数学的に許容される法則式を発見する拡張 SDS(*Extended SDS*)[9] というシステムが鷲尾等により提案されている．しかし，拡張 SDS は動的な時系列データからの法則発見を行なうことは出来ない．

このように現時点では受動的に観測される動的データから数学的に許容される法則式を発見する手法は確立されていない．本研究では，受動的に観測される動的データから数学的に許容される法則式を発見する手法の確立を目的としている．

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

2. 関連研究

2.1 法則式発見技術の概要

全ての状態変数が能動的に変更でき，かつ観測できると言う条件の下で，観測値系列からの法則発見手法としてよく知られているのは 1981 年に Langley 等により発表された BACON と BACON から派生したいくつかのシステムである．BACON をはじめとするこれらのシステムは，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関係式を組み上げる．

BACON 系に対して鷲尾等の SDS は前述のスケールタイプ制約をはじめ，単位次元解析の分野で有名な定理である *Product theorem* と *Buckinghum* Π-*theorem*[10][11] を拡張した *Extended Product Theorem*, *Extended Backinghum* Π-*theorem*[6][12]，さらに恒等式制約 (*identity constraint*)[6] を導入し，得られる法則式が数学的に許容される式であることを保証している．また，３数量組の中の任意の２数量組の間に成立する関係が無矛盾であるかを調べる三つ組試験 (*triplet test*) を行い，探索範囲をさらに減少させるという工夫もなされている．このため SDS は状態変数の数に対して高々２次ないし３次の多項式オーダーで観測データからの法則式を発見することができる．

2.2 Monte Carlo particle filter

Monte Carlo particle filter[13] は状態遷移が非線形方程式で表される対象物の状態推定に用いられるフィルタである．particle filter では N 個の系列を同時に作製し，それぞれの系列について観測値系列を観測する可能性 (尤度) を基にした重要度を付加する．今，内部状態の状態遷移を表す式を $\boldsymbol{x}_k = f(\boldsymbol{x}_{k-1}) + \mathbf{v}_k; \mathbf{v}_k \sim N(0, \Sigma_\mathrm{v})$，観測値を $\boldsymbol{y}_k$ とし，$\boldsymbol{y}_k = g(\boldsymbol{x}_k) + \mathbf{w}_k; \mathbf{w}_k \sim N(0, \Sigma_\mathrm{w})$ で表されるとする．ここで $\boldsymbol{x}$ は状態変数ベクト，$\boldsymbol{y}$ は観測変数ベクトル，$\mathbf{v}_k, \mathbf{w}_k$ はそれぞれプロセスノイズと観測ノイズベクトルである．また，各状態の重要度を表す変数を $w_k^{(i)}$ ($\sum_{i=0}^{N} w_k^{(i)} = 1$) とする．このとき，particle filter は図 1 のような操作を行って N 個の系列 (パーティクル) とそれに附随する重要度変数の値を計算する．

図 1 の中で，$\pi(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{1:k-1}^{(i)}, \boldsymbol{y}_t)$ は，i 番目のパーティクルにおいて，時刻 k における状態変数の値を決める重要度関数で

1． $\boldsymbol{x}_0^{(i)} \sim N(m_0, \Sigma_0)(i=1,2,\ldots,\mathrm{N})$ とする．

2． $1 \leq k \leq t$，$i=1,2,\ldots,\mathrm{N}$ について以下 2-1 から 2-4 をくり返す

2-1． 初期状態から時刻 $k-1$ までの状態系列 $\boldsymbol{x}_{0:k-1}^{(i)}$ と，時刻時刻 k における観測値 $\boldsymbol{y}$ から，時刻 k における内部状態 $\boldsymbol{x}_k^{(i)}$ の確率分布 $\pi\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_t\right)$ を計算する．

2-2． $\boldsymbol{x}_k^{(i)} \sim \pi\left(\boldsymbol{x}_k^{(i)} | \boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{y}_t\right)$ となるように $\boldsymbol{x}_k^{(i)}$ を 1 つサンプリングし，$\boldsymbol{x}_{1:k}^{(i))} \triangleq \left(\boldsymbol{x}_{0:k-1}^{(i)}, \boldsymbol{x}_k^{(k)}\right)$ とする．

2-3． 重要度変数を計算する．

$$w_k^{*(i)} = w_{k-1}^{(i)} p\left(\boldsymbol{y}_k \mid \boldsymbol{x}_{0:k-1}^{(i)}\right)$$

2-4． 重要度変数を正規化する．

$$w_k^{(i)} = \frac{w_k^{*(i)}}{\sum_{j=1}^{N}\left(w_k^{*(j)}\right)}$$

3． 重要度変数の値により最適な内部状態の系列 $\hat{\boldsymbol{x}}_{0:t}$ を決定する．

図 1: particle filter の動作

あり，その関数として，最適なものは，初期状態から時刻 $k-1$ までの内部状態系列が $\boldsymbol{x}_{0:k-1}^{(i)}$ で，時刻 k において観測値 $\boldsymbol{y}_k$ を観測したときの x_k の確率密度関数 $p\left(\boldsymbol{x}_k^{(i)} \mid \boldsymbol{x}_{1:k-1}^{(i)}, \boldsymbol{y}_t\right)$ である [13]．

こうして作製された N 個の系列と重要度変数から，観測値 $\boldsymbol{y}_{0:t}$ を観測する可能性が高い系列 $\hat{\boldsymbol{x}}_{0:t}$ を得る．$\hat{\boldsymbol{x}}_{0:t}$ の決定は，二乗誤差を最小とする系列を選ぶのが一般的とされているが，各時刻の重要度変数の値で重みを付けた重みつき平均値の系列や，時刻 t における重要度変数の値が最大 (事後確立が最大) となる系列を選ぶなどの方法がある．

3. 提案手法

3.1 問題設定

状態変数ベクトル $\boldsymbol{x}_k$ は連続時間系において，

$$\dot{\boldsymbol{x}} = h(\boldsymbol{x}) + \mathrm{v}_k \quad (\mathrm{v}_k \sim N(0, \Sigma_v)) \tag{1}$$

で表されるものとする．また，観測値ベクトル $\boldsymbol{y}$ は，この対象を一定時刻 Δt 間隔で観測した時系列データであるとする．ここで，観測値は適当な観測器で対象を観測することにより生成されるが，観測装置は一般に線形関数で表せることが多いので，観測過程は以下のように表す．

$$\boldsymbol{y}_k = C\boldsymbol{x}_k + \boldsymbol{d} + \mathrm{w}_k \quad (\mathrm{w}_k \sim N(0, \Sigma_\mathrm{w})) \tag{2}$$

また，計算機で連続時間系を厳密に扱うことは困難であるため，状態変数の変化は以下の式で近似する．

$$\boldsymbol{x}_k = \boldsymbol{x}_{k-1} + \dot{\boldsymbol{x}}_{k-1}\Delta t + \mathrm{v}_k \tag{3}$$

また，式 (1) の個々の微分方程式は問題を簡単にするため，“*regeme* 式”[6] に限るとする．この条件において，時刻 0 から t までの観測値 $\boldsymbol{y}_{0:t}$，観測器を表す行列 C 及びベクトル $\boldsymbol{d}$，状態遷移および観測において発生する雑音の強さ $\Sigma_\mathrm{w}, \Sigma_\mathrm{v}$，観測値 $\boldsymbol{y}$ のスケールタイプが既知であるときに対象を記述する式 (1) を発見することを目的とする．

3.2 ステップ 1：状態変数のスケールタイプを特定する

観測値 $\boldsymbol{y}$ のスケールタイプ及び，観測器を表す行列 C 及びベクトル d が既知であるので，個々の観測値 y を表す以下の式は既知となる．

$$y = \sum_{k=1}^{dim(\boldsymbol{x})} (c_k x_k) + d \tag{4}$$

ここで $dim(\boldsymbol{x})$ は $\boldsymbol{x}$ の次元数である．このとき，$\boldsymbol{y}$ のスケールタイプかとスケールタイプ制約から $\boldsymbol{x}$ のスケールタイプがある程度特定できる．スケールタイプ制約を用いると状態変数 $\boldsymbol{x}$ について，以下のような制約が得られる．

1．y が比例尺度である場合

1-1．$d=0$ かつ，$y=c_i x_i$ で表される場合
x_i のスケールタイプは比例尺度である．

1-2．$d=0$ かつ c_k の符号が全て正または 0 の場合
$c_k \neq 0$ となる k について x_k のスケールタイプは比例尺度である．

1-3．$d=0$ かつ c_k の符号が負であるものが存在する場合
$c_k \neq 0$ となる k について x_k のスケールタイプは比例尺度である．
又は，$c_k \neq 0$ である x_k のうち 2 つ以上間隔尺度である変数がある．但し，$c_k < 0$ である x_k のうち，少なくとも 1 つ以上間隔尺度である変数があり，$c_k > 0$ である x_k のうち，少なくとも 1 つ以上は間隔尺度である変数がある．

1-4．$d \neq 0$ かつ $y = c_{ik}x_i + d$ で表される場合
x_i のスケールタイプは間隔尺度である．

1-5．$d \neq 0$ で 1-4 以外の場合
状態変数のスケールタイプに制約が生じない．

2．y が間隔尺度である場合

2-1．$d=0$ かつ $y=c_i x_i$ で表される場合
x_i のスケールタイプは間隔尺度である．

2-2．$d=0$ かつ 2-1 以外の場合
$c_k \neq 0$ である k について,x_k のうち, 少なくとも 1 つの変数のスケールタイプは間隔尺度である．

2-3．$d \neq 0$ の場合
状態変数のスケールタイプに制約が生じない．

このような制約を用いて，$\boldsymbol{y}$ のスケールタイプと観測器を表す式から制約と矛盾しない $\boldsymbol{x}$ のスケールタイプの組み合わせを求める．

3.3 ステップ 2：候補式の生成

それぞれの状態変数のスケールタイプが比例尺度か間隔尺度であると仮定すると，スケールタイプ制約により状態遷移を表す微分方程式の形が限定される．つまり，ある変数 x のスケールタイプが，比例尺度と間隔尺度のどちらであってもその微分 $\frac{\mathrm{d}x}{\mathrm{d}t}$ のスケールタイプは比例尺度であるので，スケールタイプ制約により式の形が限定される．たとえば，状態変数が 2 個 (x_1, x_2) で，この 2 つの状態変数のスケールタイプが比例尺度であると仮定すると，微分方程式の形を "*regeme* 式" に限ると，x_1, x_2 の状態遷移を表す微分方程式は以下の 4 つの形しかあり得ない．

$$\frac{\mathrm{d}x_i}{\mathrm{d}t} = \begin{cases} \alpha \\ \alpha x_1^{\beta_1} \\ \alpha x_2^{\beta_2} \\ \alpha x_1^{\beta_1} x_2^{\beta_2} \end{cases} \quad (i = 1, 2)(\alpha, \beta_1, \beta_2 : \text{定数}) \tag{5}$$

このように，状態変数の数とその中の比例尺度変数と間隔尺度変数の数が分かれば "*regeme* 式" である微分方程式の形は限られるので，これらの形を予め求めて数式データバンクとして置いておき，それらの式の組み合わせを候補として次の定数の探索を行う．

3.4 ステップ 3：定数の探索

ステップ 2 において方程式の候補が決まると，式中に表れる定数をランダムに決定し，particle filter でその式を時刻 t までトラッキングし，その中で事後確立が最大となる系列について，観測器の式を適用し，模擬観測値系列 $\boldsymbol{y\prime}_{0:t}$ を得る．そして，$\boldsymbol{y\prime}_{0:t}$ と $\boldsymbol{y}_{0:t}$ の平均二乗誤差を計算し，この平均二乗誤差をその式候補の評価値として MonteCarlo 法を適用し，定数を決定する．

続けてステップ 2 とステップ 3 を繰り返し行い最終的に平均二乗誤差の小さい式の候補と定数の組み合わせを結果として出力する．

4. 予備実験

前節で述べた手法につい以下のような簡単な問題を設定して，式の発見を行うことが出来るか実験を行った．

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.5x_1 + \mathrm{w} \quad \mathrm{w} \sim N(0, 0.001^2) \tag{6}$$

$$y_1 = x_1 + \mathrm{v} \quad \mathrm{v} \sim N(0, 0.001^2) \tag{7}$$

上式の状態遷移と観測値を，観測時間間隔 $\Delta t = 0.05$，観測ステップ数 $1000(t = 0, \ldots, 999)$ の条件で作製し，particle filter のパーティクル数 $N = 500$ として前節で述べた方法を適用して得られた結果のうち観測値の平均二乗誤差の小さい順に 5 つを以下の表 1 に示す．

表 1: 予備実験結果

順位	候補式	平均二乗誤差
1	$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.92x_1$	0.00086
2	$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.48x_1$	0.00127
3	$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.84x_1$	0.00139
4	$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.72x_1$	0.00154
5	$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -0.81x_1$	0.00156

表 1 より，提案手法で得られた式は元の式の形を正しく再現しており，特に 2 番目の式は真の式に近いという結果が得られた．

このように提案手法は観測値系列と観測器を表す式が分かれば，対象を表す法則微分方程式を発見できる可能性を有していることが分かる．

5. 結論

前節の予備実験の結果から，簡単な問題に対しては提案手法が適用すれば対象の観測値系列からその対象を支配する法則微分方程式の発見が分かった．現在，より複雑な問題に対して追加の実験を行っているが，現段階では満足すべき結果が得られていない．その理由として，式の候補の生成をランダムに行っていることと，定数の同定を Monte Carlo 法で行っていることが挙げられる．前者に関しては，問題が複雑になるに従って，可能な式の候補の数が指数的に増大するため，ランダム探索では取りこぼしが多すぎると考えられる．また，後者についても，式中に現れる係数のうち，べき係数が少しでも変わると方程式の挙動が大きく変わるので，少なくともべき係数に関しては Monte Carlo 法が適していないことが考えられる．これらの知見を踏まえた上で，より効率的な探索方法を今後の研究の課題としていきたい．

参考文献

[1] P.W.Langley, "*Data-Drivendiscovery of phisical laws*", Cognitive Science, Vol. 5, pp.31-54,1981

[2] P.W.Langley, H.A.Simon, G.Bradshaw and J.M.Zytkow, "*Scientific Discovery; Computational Explorations of the Creative Process*", Cambridge, Massachusetts: MIT Press, 1987

[3] B.Koehn and J.M.Zytkow, " *Experimenting and theorizing in theory formation*", Proc. of the International Symposium on Methodologies for Intelligent Systems, pp.296-307, 1986

[4] B.C.Falkenhainer and R.S.Michalski, "*Intergrating Quantitative and Quantitative Discovery; The ABACUS System*", Machine Learning, pp.367-401,Kluwer Academic Publishers, 1986

[5] R.D.Luce, "*On the Possible Psychological Laws* ", The Psychological Review, Vol.66, No.2, pp.81-95, 1959

[6] T.Washio and H.Motoda, "*Discovering Admissible Models of Complex Systems Based on Cale Types and Identity Constraints* ", Proc. of the 15th International Joint Conference on Artificial Intelligence (IJCAI), Vol.2, pp.810-817,1997

[7] S.S.Stevence, "*On the Theory of Scales of Measurement* ", *SCIENCE*, Vol.103, No.2684, pp.677-680, 1946

[8] S.Dzeroski and L.Todorovski, "*Discovering Dynamics; From Inductive Logic Programing to Machine Discovery* ", Journal of Intelligent Information Systems, pp.1-20, Kluwer Academic Publishers, 1994

[9] T.Washi, H.Motoda and Y.Niwa, "*Discovering Admissible Model Equations form Observed Data Based On Scale-types and Identity Constraints* ", Proc. of the 16th International Joint Conference on Artificial Intelligence (IJCAI), Vol.2, pp.772-779, 1999

[10] P.W.Bridgeman, "*Dimensional Analysis* ", Yale University Press, New Haven, CT, 1922

[11] E.Buckinghum, "*On Physically similar systems; Illustrations of the use of dimensional equations* ", Physical Review, Vol.IV, No.4, pp.345-376, 1914

[12] T.Washio and H.Motoda, "*Extensions of Dimensional Analysis for Scale-type and its Application to Discovery of Admissible Models of Complex Processes* ", Working Papers of Similarity Workshop'99; The Second International Workshop on Similarity Models, University Stuttgart, Germany, 1999

[13] A.Doucet, S.Godsill and C.Andriew,"*On sequencial Monte Carlo sampling methods for Bayesian filtering* ", Statistics and Computiong, 10, pp.197-208, 2000