力学系のファジィモデルによる知識発見手法の一提案

3U－10

竹内一郎 古橋 武
(名古屋大学大学院工学研究科)

1　はじめに

モデルをエキスパートシステムやデータマイニング等に応用する際には知識の導入や解析の容易性などが要求される。ファジィモデリングは、記号的知識をモデルに導入することが容易な点や領域ごとの知識が構築後のモデルに表現される点などからこれに適している。本稿では、時系列を扱うモデルにファジィモデリングの概念を適用し、知識の導入や解析が容易となる時系列モデルの構築法を提案する。

本稿で考察するモデリングは、「時系列データと記号列で表現されるヒューリスティックな知識からモデルを構築すること」と定式化される。提案する手法にでは、ヒューリスティックな記号列が記号力学系から生ずるとの仮定に基づき、与えられた記号列を時系列データの背後にある力学系の状態空間の表現に対応させる。

提案手法によるモデルの構造は、記号力学系の各離散状態に対応するサブモデルの集合により表現され、サブモデルの機能を定めることにより、線形か非線形か、自励系か非自励系かのいづれの力学系も表現可能である。サブモデル間の遷移により時系列を表現するモデルは数多く提案されており[1] [2] 、このような構造が時系列の表現に適していることを示唆している。しかし、これらの従来モデルは、状態遷移を陽に扱っておらず、これを力学系の知識表現に用いる視点はみられない。

本稿では、最も解析が困難である非線形非自励系を考察する。2章でモデルのクラスを定め、3章でモデルの同定について説明し、4章でまとめる。

2　モデルのクラス

本章の目的は、時系列データ$\boldsymbol{Z}$ とヒューリスティックな記号列$\boldsymbol{S}$ が与えられたときに、$\boldsymbol{Z}$ の内部表現を$\boldsymbol{S}$ と関連させて表現できるモデルのクラスを見付けることである。本稿では、状態がファジィに表現される記号力学系のモデルによりこれを実現する。まず、記号力学系による時系列の表現について説明し、続いて、その系の離散状態をファジィ化させることでモデルのクラスを導く。

記号力学系は、離散有限個の状態とそれら状態間の遷移関係により定義される。記号力学系の各状態に関数を定義すると、この記号力学系における任意の状態遷移に対して時系列データが生成される。図1にこのような系による時系列の表現を示す。系は各状態 s_i における2つの関数$\boldsymbol{f}_i \in \boldsymbol{F}$ 、$g_i \in \boldsymbol{G}$ によって定義される。m 番目の状態 $s(m) = s_i$ にある系への入出力を$\boldsymbol{u}_m$、$\boldsymbol{y}_m$ とすると、$\boldsymbol{f}_i$ 、g_i は次のように機能する。

$$\boldsymbol{y}_m = \boldsymbol{f}_i(\boldsymbol{u}_m) \tag{1}$$

$$\Delta_m s(m) = g_i(\boldsymbol{u}_m) \tag{2}$$

ここで、Δ はその添字に関する前進差分を表す演算子である。なお、ここでは、状態遷移にマルコフ性を仮定している。

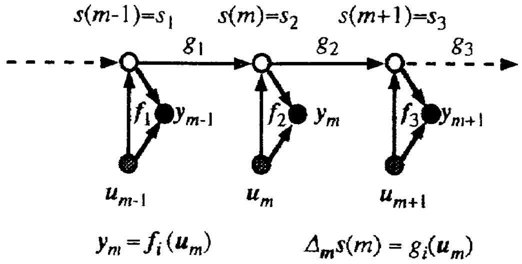

図 1: Symbolic Dynamics

2つの関数群$\boldsymbol{F}$ 、$\boldsymbol{G}$ によって定義される系に対して、各時刻 t の状態$\boldsymbol{x}(t)$ を各離散状態 s_i のメンバーシップのグレード $x_i(t)$ の集合によって表現することを考える。このとき、(1)(2)式は次のように変わる。

$$\boldsymbol{y}_t = \sum_i x_i(t)\boldsymbol{f}_i(\boldsymbol{u}_t) \tag{3}$$

$$\Delta_t \boldsymbol{x}_t = \sum_i x_i(t) g_i(\boldsymbol{u}_t) \tag{4}$$

ここで、離散状態の個数を N とすると、$\boldsymbol{x}(t)$ は $[0,1]^N$ 上の点群 $\{(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,1)\}$ を端点とする凸多面体$\boldsymbol{P}$ 上を動き、関数群$\boldsymbol{G}$ は各時刻 t の入力$\boldsymbol{u}_t$ により、ベクトル場$\boldsymbol{P} \to \boldsymbol{P}$ を定める役割を担う。図2は $N = 3$ の場合の凸多面体$\boldsymbol{P}$ を表しており、$\boldsymbol{P}$ の各端点に離散状態 s_i が対応している。図の黒丸は系のダイナミクス$\boldsymbol{X} = \{\boldsymbol{x}(1),\boldsymbol{x}(2),\cdots,\boldsymbol{x}(T)\}$ を表している。ここで、$\boldsymbol{S}: s_1 \to s_2$ を系のダイナミクス

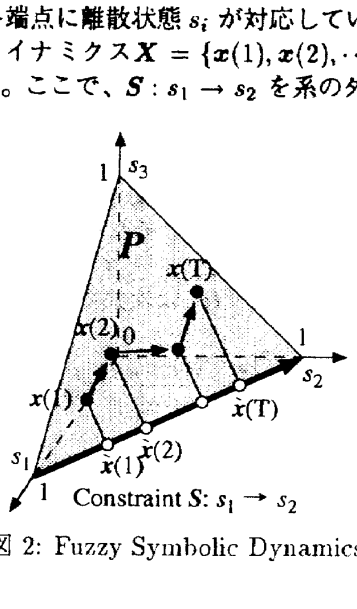

図 2: Fuzzy Symbolic Dynamics

A Proposal of Knowledge Discovery Method by Dynamical Fuzzy Modeling

Ichiro Takeuchi, Takeshi Furuhashi
Nagoya University Graduate School of Engineering, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

に関するヒューリスティックな知識とすると、このダイナミクスは、時間と共に s_1 から s_2 方向へ両者を結ぶ線分上を動く運動で表現される。$\boldsymbol{S}$ をダイナミクスを同定する際の制約条件と考えると、制約$\boldsymbol{S}$ は$\boldsymbol{X}$ をこの線分上に表現することで満たされる。すなわち、$\boldsymbol{X}$ に制約$\boldsymbol{S}: s_1 \rightarrow s_2$ が加わると、ダイナミクスは$\boldsymbol{X}$ を線分 s_1s_2 上へ射影した$\hat{\boldsymbol{X}} = \{\hat{x}(1), \hat{x}(2), \cdots, \hat{x}(T)\}$ で表される。

3 モデルの同定

モデリングの際に与えられる情報は L 個の時系列データ$\boldsymbol{Z}^{*1}, \boldsymbol{Z}^{*2}, \cdots, \boldsymbol{Z}^{*L}$ とそれぞれに対応するヒューリスティックな記号列$\boldsymbol{S}^1, \boldsymbol{S}^2, \cdots, \boldsymbol{S}^L$ である。$\boldsymbol{Z}^{*l}$ は長さ $T(l)$ の入力系列$\boldsymbol{U}^l = \{\boldsymbol{u}_1^l, \boldsymbol{u}_2^l, \cdots, \boldsymbol{u}_{T(l)}^l\}$ と出力系列$\boldsymbol{Y}^{*l} = \{\boldsymbol{y}_1^{*l}, \boldsymbol{y}_2^{*l}, \cdots, \boldsymbol{y}_{T(l)}^{*l}\}$ から成り、$\boldsymbol{S}^l$ は長さ $M(l)$ の記号列 $\{S_1^l, S_2^l, \cdots, S_{M(l)}^l\}$ である。

与えられる L 個の記号列$\boldsymbol{S}^1, \boldsymbol{S}^2, \cdots, \boldsymbol{S}^L$ の全種類の記号を要素とする集合を$\boldsymbol{s}$ とする。$\boldsymbol{s}$ の各要素を離散状態とする記号力学系を考えると、すべての記号列はこの記号力学系の状態遷移により表現できる。

さて、前章で考察したように、この記号力学系をファジィ化し、関数群$\boldsymbol{F}$、$\boldsymbol{G}$ を定めることでモデルの構造が定まる。このモデルに入力系列$\boldsymbol{U}^l = \{\boldsymbol{u}_1^l, \boldsymbol{u}_2^l, \cdots, \boldsymbol{u}_{T(l)}^l\}$ が与えられると、モデルは関数群$\boldsymbol{F}$、$\boldsymbol{G}$ により、状態遷移系列$\boldsymbol{X}^l = \{\boldsymbol{x}^l(1), \boldsymbol{x}^l(2), \cdots, \boldsymbol{x}^l(T(l))\}$ および出力系列$\boldsymbol{Y}^l = \{\boldsymbol{y}^l(1), \boldsymbol{y}^l(2), \cdots, \boldsymbol{y}^l(T(l))\}$ を導く。ここで、モデルの同定とは$\boldsymbol{X}^l$ が$\boldsymbol{S}^l$ に従うという制約のもとで、$\boldsymbol{Y}^l$ が$\boldsymbol{Y}^{*l}$ をできるかぎり近似するように、関数群$\boldsymbol{F}$、$\boldsymbol{G}$ のパラメータ調整を行うことを意味する。

パラメータ調整にはEMアルゴリズムを用いる。簡単には、$\boldsymbol{F}$ を固定して$\boldsymbol{G}$ を適応させる E-step と$\boldsymbol{G}$ を固定して$\boldsymbol{F}$ を適応させる M-step とを繰り返すことにより実現される。vステップ後の両関数群を$\boldsymbol{F}^v$、$\boldsymbol{G}^v$、また$\boldsymbol{G}^v$ により同定された状態遷移系列を$\boldsymbol{X}^{l,v}$ とする。以下に両ステップの説明を行う。

E-step: E-step の目的は、固定された$\boldsymbol{F}^v$ と$\boldsymbol{S}^l$ の制約のもとで、最良の状態遷移を実現できるように$\boldsymbol{G}^v$ を適応させることである。そのため、まず、目標とする状態遷移系列$\hat{\boldsymbol{X}}^{l,v}$ を求める。

前章の図2のように、$\boldsymbol{X}^{l,v} = \{\boldsymbol{x}^{l,v}(1), \boldsymbol{x}^{l,v}(2), \cdots, \boldsymbol{x}^{l,v}(T(l))\}$ は凸多面体$\boldsymbol{P}$ 上の点列で表される。これらの点を線分 $S_m^l S_{m+1}^l$ 上に射影することで、$\boldsymbol{S}^l$ の制約が付加される。S_m^l、S_{m+1}^l それぞれに対応する端点の位置ベクトルを$\boldsymbol{e}_m^l$、$\boldsymbol{e}_{m+1}^l$ とすると$\boldsymbol{x}^{l,v}(t)$ の線分 $S_m^l S_{m+1}^l$ 上への射影$\hat{\boldsymbol{x}}^{l,v}(t)$ は次式で表される。

$$\hat{\boldsymbol{x}}^{l,v}(t) = \frac{(\boldsymbol{e}_{m+1}^l - \boldsymbol{e}_m^l)^T(\boldsymbol{x}^{l,v}(t) - \boldsymbol{e}_m^l)}{(\boldsymbol{e}_{m+1}^l - \boldsymbol{e}_m^l)^T(\boldsymbol{e}_{m+1}^l - \boldsymbol{e}_m^l)}(\boldsymbol{e}_{m+1}^l - \boldsymbol{e}_m^l) + \boldsymbol{e}_m^l \quad (5)$$

ここで、T は転置を表す。その後、各点$\hat{\boldsymbol{x}}^{l,v}(t)$ に関して、次式のように、線分 $S_m^l S_{m+1}^l$ 上の両側を探索し、近似が進む方向へ変位させて$\hat{\boldsymbol{x}}^{l,v}(t)$ を得る。

$$\hat{\boldsymbol{x}}^{l,v}(t) = \hat{\boldsymbol{x}}^{l,v}(t) + \alpha \frac{\partial E1_t^{l,v}(\hat{\boldsymbol{x}}^{l,v}(t))}{\partial \hat{x}_{m+1}^{l,v}(t)}(\boldsymbol{e}_{m+1}^l - \boldsymbol{e}_m^l) \quad (6)$$

ここで、α は学習係数、$E1_t^{l,v}(\boldsymbol{x})$ は入力$\boldsymbol{u}_t^l$ に対する状態が$\boldsymbol{x}$ で表されるモデルの出力$\boldsymbol{y}_t^{l,v}$ と$\boldsymbol{y}_t^{*l}$ との2乗誤差で、次式のように表される。

$$E1_t^{l,v}(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{y}_t^{*l} - \boldsymbol{y}_t^{l,v}|^2 \quad (7)$$

$$\boldsymbol{y}_t^{l,v} = \sum_i x_i \boldsymbol{f}_i^v(\boldsymbol{u}_t^l) \quad (8)$$

また、x_m は S_m に対応する離散状態 $s(m)$ のメンバーシップのグレードを表す。

(6) 式より計算した$\hat{\boldsymbol{X}}^{l,v}$ を実現するように関数群$\boldsymbol{G}^v$ を学習し、$\boldsymbol{G}^{v+1}$ を得る。次式で表される2乗誤差

$$E2_t^l = \frac{1}{2}|\Delta_t \hat{\boldsymbol{x}}^{l,v}(t) - \sum_i \hat{x}_i^{l,v}(t) \cdot \boldsymbol{g}_i^v(\boldsymbol{u}_t^l)|^2 \quad (9)$$

に対し、最急降下法により関数$\boldsymbol{g}_i^v$ の各パラメータを更新し、$\boldsymbol{g}_i^{v+1}$ を得る。なお、関数群$\boldsymbol{G}$ は各パラメータに対して微分可能な関数とする。

M-step: M-step の目的は固定された$\boldsymbol{G}^{v+1}$ によって状態遷移系列$\boldsymbol{X}^{l,v+1}$ が決まる時、近似が進むように$\boldsymbol{F}^v$ を適応させることである。

まず、$\boldsymbol{G}^{v+1}$ と$\boldsymbol{x}_t^{l,v}$ を用いて次式のように$\boldsymbol{X}^{l,v+1}$ が求める。

$$\boldsymbol{x}^{l,v+1}(1) = \boldsymbol{x}^{l,v}(1)$$

$$\boldsymbol{x}^{l,v+1}(t+1) = \boldsymbol{x}^{l,v}(t) + \sum_i x_i^{l,v}(t) \boldsymbol{g}_i^{v+1}(\boldsymbol{u}_t^l) \quad (10)$$

この状態遷移系列$\boldsymbol{X}^{l,v+1}$ に従うとき、入力$\boldsymbol{u}_t^l$ に対するモデルの出力$\boldsymbol{y}_t^{l,v+1}$ と$\boldsymbol{y}_t^{*l}$ との2乗誤差 $E1_t^{l,v+1}(\boldsymbol{x}_t^{l,v+1})$ に対し、最急降下法により関数$\boldsymbol{f}_i^v$ の各パラメータを更新し、$\boldsymbol{f}_i^{v+1}$ を得る。なお、関数群$\boldsymbol{F}$ は各パラメータに対して微分可能な関数とする。

4 まとめ

本稿では、時系列データと記号列で与えられるヒューリスティックな知識をもとにしたモデリング手法を提案した。発表時に、モデルの検証のための数値実験を紹介する。

参考文献

[1] L. R. Rabiner and B. H. Juang. An introduction to hidden markov model. *IEEE ASSP Magazine*, pp. 4-16, 1996.

[2] R. E. Quandt. A new approach to estimating switching regressions. *Journal of the American Statistical Association*, Vol. 67, pp. 306-310, 1972.