決定性有限オートマトンのオンライン同定

渡辺　浩司
(メディア情報文化学科)

本報告では未知システムとして有限オートマトン (FA) を与え, 観測される入出力データから決定性 FA (DFA) を実現する DFA のオンライン同定について検討する.

FA の同定はオンライン同定とオフライン同定の二つに分類される. オフライン同定とは同定対象の FA が必要に応じて初期状態にリセットできるものであり, 同定に必要な入出力データは繰返し実験によりすべて得ることができる. 文献 [1] の無限ハンケル行列からの最小 DFA の実現はオフライン同定である. これに対し今回行なうオンライン同定は初期状態へのリセットが不可能であるという制約が同定対象に対して加えられている. そのためオフライン同定で用いる繰返し実験を行なうことができない. そこで今回は同定対象の FA の受理状態は一つであるという仮定をし, 受理状態を初期状態とみなすことにより繰返し実験を可能にしている.

【キーワード　有限オートマトン　 状態空間モデル　 オンライン同定】

1　まえがき

現代制御理論の分野において, ある未知システムを外部から観測してその入出力データを収集し, それを元に未知システムの伝達関数, 微分方程式, 状態空間モデルといった数学モデルを決定することを同定と呼ぶ.

本報告では未知システムとして有限オートマトン (FA) を与え, 観測される入出力データから決定性 FA (DFA) を実現する DFA のオンライン同定について検討する.

FA の同定はオンライン同定とオフライン同定の二つに分類される. オフライン同定とは同定対象の FA が必要に応じて初期状態にリセットできるものであり, 同定に必要な入出力データは繰返し実験によりすべて得ることができる. 先に報告した無限ハンケル行列からの最小 DFA の実現 [1] はオフライン同定である. これに対し今回行なうオンライン同定は初期状態へのリセットが不可能であるという制約が同定対象に対して加えられている. そのためオフライン同定で用いる繰返し実験を行なうことができない. そこで今回は同定対象の FA の受理状態は一つであるという仮定をし, 受理状態を初期状態とみなすことにより繰返し実験を可能にしている.

なお, FA は離散時間動的システムとみなすことができ, ブール半環 $B(=(B,+,\cdot), B=\{0,1\})$ 上の状態空間モデルとして表現できるので, ここで同定されるのはシステムパラメータ $(\{A_k\}, \boldsymbol{c}, \boldsymbol{x}_0)$ である.

2　有限オートマトンの状態空間モデル

2.1　状態空間モデル

n 状態, m 入力の FA は次のような B 上の状態空間モデルで表現できる.

$$\begin{cases} \boldsymbol{x}(t+1) = \displaystyle\sum_{k=1}^{m} u_k(t) A_k \boldsymbol{x}(t) \\ y(t) = \boldsymbol{c}^t \boldsymbol{x}(t) \end{cases} \tag{1}$$

$\boldsymbol{x}(\in B^n)$ は状態, $y(\in B)$ は出力, $u_k(\in B)$, $A_k(\in B^{n\times n})$ は記号 a_k $(\in \Sigma:$ 記号集合$)$ の入力とそれによる状態推移, $\boldsymbol{c}(\in B^n)$ は出力ベクトルを表す. 初期状態を $\boldsymbol{x}_0$ とし, $\boldsymbol{x}(0) = \boldsymbol{x}_0$ である. 出力は $y(t) = 1/0$ となり, それぞれ受理/非受理を表す. なお, 式 (1) に対する FA のパラメータ表現を $(\{A_k\}, \boldsymbol{c}, \boldsymbol{x}_0)$ $(k = 1, \cdots, m)$ と表す.

2.2　可到達性および可観測性

入力記号列 $w = a_{k_{t-1}} a_{k_{t-2}} \cdots a_{k_0}$ $(\in \Sigma^*)$ を FA に入力するとき, 右の文字から入力されるものとし, w に対応する状態推移行列 $A(w)$ を次のように定義する.

$$A(w) = A_{k_{t-1}} A_{k_{t-2}} \cdots A_{k_1} A_{k_0} \tag{2}$$

ただし, a_{k_t} $(\in \Sigma)$ は時刻 t における入力記号, A_{k_t} $(\in \{A_1, \cdots, A_m\})$ は a_{k_t} に対応する状態推移行列である.

次の行列 $R_{n,\infty}$ を可到達行列という.

$$\begin{aligned} R_{n,\infty} = [\ & \boldsymbol{x}_0, A_1\boldsymbol{x}_0, \cdots, A_m\boldsymbol{x}_0, A_1A_1\boldsymbol{x}_0, \\ & \cdots, A_mA_1\boldsymbol{x}_0, \cdots, A(w)\boldsymbol{x}_0, \cdots]\ \ (w \in \Sigma^*) \end{aligned} \tag{3}$$

この $R_{n,\infty}$ の i 行目は状態 q_i に対応し, i 行目が零行ベクトルのとき, q_i は不可到達であるといい, 初期状態からその状態へ推移する入力記号列は存在しない. また $R_{n,\infty}$ に零行ベクトルが存在しないとき $(\{A_k\}, \boldsymbol{x}_0)$ は完全可到達であるという.

次の行列

$$O_{\infty,n} = \begin{pmatrix} \boldsymbol{c}^t \\ \vdots \\ \boldsymbol{c}^t A(w) \\ \vdots \end{pmatrix} \quad (w \in \Sigma^*) \tag{4}$$

を可観測行列という．この可観測行列の i 列目は状態 q_i に対応し i 列目が零列ベクトルであるとき, q_i は不可観測であるといい, その状態から受理状態へ推移する入力記号列は存在しない．また同じ列が存在する時, それらに対応する状態は出力側から見て等価な状態であり, 識別不可能と呼ぶ．また $O_{\infty,n}$ に零列ベクトルが存在しないとき $(\{A_k\}, \boldsymbol{c})$ は完全可観測であるという．

2.3 特性応答とハンケル行列

初期状態が $\boldsymbol{x}_0$ で入力記号列が w のとき, 状態空間モデル (1) の一般解は次のように書ける．

$$\boldsymbol{x}(t) = A(w)\boldsymbol{x}_0 \quad , \quad y(t) = \boldsymbol{c}^t A(w)\boldsymbol{x}_0 \tag{5}$$

次に状態空間モデル (1) の入出力関係を表す特性応答を次のように定義する．

$$\{h(\varepsilon), h(a_1), \cdots, h(w), \cdots\} \quad ({}^\forall w \in \Sigma^*) \tag{6}$$

ここで ε は空記号列とし, $A(\varepsilon)$ は n 次単位行列とする．$h(w)$ は入力記号列が w のときの FA の出力 (応答) であり, 次式のように書ける．

$$h(w) = \boldsymbol{c}^t A(w)\boldsymbol{x}_0 \tag{7}$$

この $h(w)$ を用いて, ハンケル行列と呼ばれる無限行列 $\mathcal{H}_{\infty,\infty}$ を次のように構成する．

$$\mathcal{H}_{\infty,\infty} = \begin{array}{c} \\ \varepsilon \\ \vdots \\ r \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} \varepsilon & \cdots & v & \cdots \end{array} \\ \begin{pmatrix} h(\varepsilon) & \cdots & h(v) & \cdots \\ \vdots & & \vdots & \\ h(r) & \cdots & h(rv) & \cdots \\ \vdots & & \vdots & \ddots \end{pmatrix} \end{array} \tag{8}$$

ここで r $(\in \Sigma^*)$ は行ラベル, v $(\in \Sigma^*)$ は列ラベルであり, いずれも辞書順に並べられる．

ハンケル行列 $\mathcal{H}_{\infty,\infty}$ は可観測行列 $O_{\infty,n}$ と可到達行列 $R_{n,\infty}$ の積に分解できる．

$$\mathcal{H}_{\infty,\infty} = O_{\infty,n} R_{n,\infty} \tag{9}$$

ある DFA が完全可到達かつ完全可観測で全状態が識別可能 (完全可識別) であるとき, すなわち, $R_{n,\infty}$ に n 個の単位ベクトルがすべて存在し, $O_{\infty,n}$ の n 本の列ベクトルが全て非零で互いに異なるとき, その DFA の状態空間モデルは最小次元になっている．このとき $\mathcal{H}_{\infty,\infty}$ はその列ベクトルが $O_{\infty,n}$ の n 個の異なる非零列ベクトルと零ベクトルだけからなっている．よって, あるハンケル行列が得られたとき, その列ベクトルのうち異なる非零列ベクトルの数 $\overline{n}$ が最小 DFA の状態数を表していることが分かる．

3　決定性有限オートマトンのオンライン同定アルゴリズム

本章では DFA のオンライン同定のアルゴリズムを述べる. まず同定対象となる DFA に対して以下のことを仮定する.

1. 受理状態は一つ
2. 状態数の上限は既知
3. 完全定義の死状態を持たない最小 DFA

オンライン同定では, 同定対象を初期状態にリセットできないためオフライン同定における繰返し実験を行なえず, 同定に必要なデータを得ることが困難である. そこで受理状態を初期状態とみなすことにし, 受理状態に到達したときシステムは初期状態にリセットされたものとする. 仮定 1. より受理状態は一つなので初期状態を特定できることになる.

入力に関しては基本的にランダムに与えられ, 必要に応じて観測者が適当な入力を設計できるものとする. ただしランダムな入力系列は状態推移図におけるすべての推移を少なくとも一回は通り (励起し), 状態推移関数を決定できる有限系列 (励起系列) であるとする.

入力系列に対するこの条件と仮定 2. より入出力応答を十分に観測することにより同定に必要なデータはすべて得られることになる.

仮定 3. については不完全定義の DFA や死状態を持つ DFA が同定対象となった場合は同定自体が不可能になるために加えたものである.

以下で DFA のオンライン同定のアルゴリズムを例題を用いて解説する. なおハンケル行列からの DFA の実現には先に報告した DFA の最小部分実現のアルゴリズム [1] を用いる.

<u>例題</u>

図 1 の DFA M_1 を同定対象とする. 状態数の上限は 2 とする. 観測は M_1 がランダムな入力系列により受理状態に到達した時に開始するため, 図 1 のように初期状態 = 受理状態となっているモデルを対象としている.

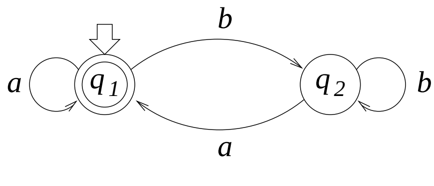

図 1　同定対象 M_1

1. ランダムな入力系列 $w = aba$ が加えられた時点で得られる情報は以下の通りである.

$$\text{受理記号列} \quad : \quad \varepsilon\ ,\ a\ ,\ ab$$
$$\text{非受理記号列} \quad : \quad b$$

初期状態 = 受理状態なので ε は受理となる．次に入力系列の最初の記号である a が入力されると受理となる．つまり a は受理記号 (列) となる．ここで一旦 M_1 はリセットされたことになり次に続く入力 ab は初期状態からの入力となる．まず b が入力されるが非受理であり, 続けて a が加えられ受理となる．結局上のような情報が得られることになる．
この情報よりハンケル行列を構成する．

```
  |e a b a b a b
  |      a a b b
-----------------
 e|1 1 0 . . 1 .
 a|1 . 1 . . . .
 b|0 . . . . . .
```

M_1 の状態数の上限が 2 と既知なので, ハンケル行列には異なる列は最大 2 本しか存在しない．つまりこの場合は ε 列と b 列が異なる列となり, a, ab 列は ε 列に等しいことがわかる．このようにして未定部分を可能な限り埋めていくと次のハンケル行列が得られる．

```
  |e a b a b a b
  |      a a b b
-----------------
 e|1 1 0 1 0 1 .
 a|1 1 1 1 1 1 .
 b|0 0 . 0 . 0 .
```

このハンケル行列からは図 2 の DFA が実現される．

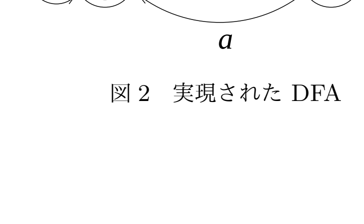

図 2　実現された DFA

先ほど得られたハンケル行列では状態 q_2 からの入力記号 b に対する推移を求められない. これはハンケル行列の bb 列がすべて未定であるため ε 列に等しいか b 列に等しいかが決められないためである. そこでこの推移を決定するために入力 $w' = bb$ を設計し M_1 に加える. 入力 $w = aba$ が加えられた時点で同定対象の DFA は受理状態 (= 初期状態) に推移しているためそのまま入力することができる.

2. $w' = bb$ が入力された時点での出力は 0, つまり非受理なのでハンケル行列は次のようになる.

```
  |e a b a b a b
  |      a a b b
-----------------
 e|1 1 0 1 0 1 0
 a|1 1 1 1 1 1 .
 b|0 0 0 0 . 0 .
```

bb 列は b 列に等しいことがわかり, 最終的に次のハンケル行列が得られる.

```
  |e a b a b a b
  |      a a b b
-----------------
 e|1 1 0 1 0 1 0
 a|1 1 1 1 1 1 1
 b|0 0 0 0 0 0 0
```

このハンケル行列から DFA を実現すると同定対象である図 1 の DFA M_1 が実現される.

4　ホーミング系列による状態同定

3 章で述べた DFA のオンライン同定のアルゴリズムでは同定の対象となる FA に対し受理状態は一つであるという制約を課した. これは一つの状態を特定しなければオフライン同定における繰返し実験を行なうことができないためであるがこの制約は一般的とはいえない. しかしこの制約を外すためには同定対象の FA の現在の状態が特定できなければならない. ただし, すべての状態を特定する必要はなく, 一つの状態が特定できればその状態を初期状態とみなせるため繰返し実験が実施できる.

ここでは順序機械における状態同定の手法であるホーミング系列 [2] について述べ, FA のオンライン同定での利用について検討する.

順序機械のホーミング系列は次のように定義される.

ホーミング系列　順序機械の状態集合 Q の部分集合 Q_I のいずれかの状態から出発するものとし, 入力系列 $\widetilde{\boldsymbol{x}}_H$ に対する出力応答を観測することによって推移した後の最終状態を一義的に決定できるとき, この $\widetilde{\boldsymbol{x}}_H$ を (Q_I に関する) ホーミング系列という.

次の状態推移表で表される順序機械 S のホーミング系列を求めると $\widetilde{\boldsymbol{x}}_H = bbabb$ となる. これは判定木と呼ばれる木を表 1. から構成することにより得られるがここではその構成法は省略する.

表 1　順序機械 S

	δ		ω	
	a	b	a	b
q_1	q_2	q_6	0	0
q_2	q_3	q_2	0	1
q_3	q_4	q_3	0	1
q_4	q_5	q_1	0	1
q_5	q_3	q_6	0	1
q_6	q_4	q_5	0	1

ただし δ は推移関数, ω は出力関数を表すものとし, q_i $(i = 1, \cdots, 6)$ は状態を表す.

S のすべての状態について $\widetilde{\boldsymbol{x}}_H = bbabb$ を入力としたときの出力系列と最終状態を求めると次のようになる.

表 2　ホーミング系列 $bbabb$ による出力系列と最終状態

初期状態	出力系列	最終状態
q_1	01011	q_3
q_2	11011	q_3
q_3	11010	q_6
q_4	10010	q_6
q_5	11011	q_3
q_6	11010	q_6

以上より, たとえ初期状態がわからなくても $bbabb$ を加えたときの出力系列が

$$\left.\begin{array}{l}01011\\11011\end{array}\right\}\text{のときの最終状態は } q_3$$

$$\left.\begin{array}{l}11010\\10010\end{array}\right\}\text{のときの最終状態は } q_6$$

と決定できる．

このようにホーミング系列がわかれば状態の同定が可能となる．ここではホーミング系列の求め方については省略したが, ホーミング系列は順序機械の推移関数等がすべてわかっている場合にのみ求めることができる．

FA も順序機械の一種なので推移関数等がすべて決まっているならばホーミング系列を求めることができる．また, ハンケル行列からでもホーミング系列を求めることができる．次の DFA M_2 のホーミング系列をハンケル行列から求めてみる．

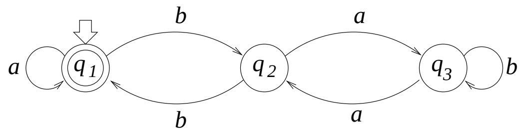

図 3　3 状態 DFA M_2

M_2 のハンケル行列は次のようになる．

```
  |e a b a b a b
  |      a a b b
-----------------
 e|1 1 0 1 0 0 1
 a|1 1 0 1 0 0 1
 b|0 0 1 0 1 0 0
aa|1 1 0 1 0 0 1
ba|0 0 0 0 0 1 0
ab|0 0 1 0 1 0 0
bb|1 1 0 1 0 0 1
```

異なる状態は ε, b, ab であり, この 3 状態を出力系列により識別できる入力がホーミング系列となる．まず a 行に注目する．a 行 ε 列は 1 であり b, ab 列は 0 なので入力 a を加えて 1 を出力したときは推移後の状態が ε であると同定できる．次に ba 行に注目すると今度は b 列のみが 1 となっている．つまり, ba を加えて 1 を出力した場合, 現在の状態が ba であることがわかる．入力列 ba に対する出力と推移後の状態は次の表 3. のようになる．

表 3　入力列 ba による出力系列と推移後の状態

推移前	出力系列	推移後
q_1	01	q_2
q_2	00	q_3
q_3	10	q_1

入力列 ba による出力を観測することにより推移後の状態を同定できるので ba はホーミング系列であることがわかる.

しかし, オンライン学習では当然ハンケル行列は未知である. 状態を同定するためにはハンケル行列が必要となるが, ハンケル行列を構成するためには状態の同定が必要となる. 受理状態を一つに限定するという条件を取り除くにはこの問題を解決する必要がある.

5　まとめ

有限オートマトン (FA) のオンライン同定について検討した. 今回は同定対象である FA の受理状態が一つであるという仮定を加え, 受理状態を初期状態とみなして繰返し実験を行なった. この受理状態が一つという仮定を外すためには何らかの方法を用いて現在の状態を特定することが必要となる. これは順序機械の最終状態を同定するホーミング系列を利用することにより可能であると考えられる. ただしホーミング系列は本来順序機械の推移関数等がすべて既知である場合に得られるものであるため, FA のオンライン同定では容易には求められない. このホーミング系列をいかにして求めるかが今後の課題となる.

参考文献

[1]　渡辺, 猪飼, 福永, “Ho-Kalman アルゴリズムによる決定性有限オートマトンの最小実現”, 信学技報, COMP97-52(1997).

[2]　当麻喜弘, “順序回路論”, 昭晃堂.

Online Identifications for Finite Automata

Koji WATANABE

Regarding finite automata (FAs) as discrete time dynamical systems, they can be represented as state space models over $B(=\{0,1\})$ similar to the representation method of linear systems over the real numbers (R) in the field of dynamical systems and controls. Based on this representation, we propose an online state identification method for deterministic FAs.

【keyword　finite automata　state space models　online identification】